Developing a psychological model of end-users' experience with news Web sites

Gabor Aranyi

A thesis submitted in partial fulfilment of
the requirements of Teesside University
for the degree of Doctor of Philosophy

May 2012

School of Social Sciences and Law
Teesside University
Middlesbrough, United Kingdom

DECLARATION

While registered as a candidate for the degree of Doctor of Philosophy the author has not been registered for any other award with any other university or institution. No part of the material in this thesis has been submitted for any degree or other qualification at any other institution by the author or, to the best of his knowledge and belief, by any other person. The thesis describes the author's original work.

ABSTRACT

The primary aim of the research project presented in this thesis was to develop and test a comprehensive psychological model of interaction experience with news Web sites. Although news media have been publishing on the Web increasingly since the second half of the 1990s and news sites have become a favoured source of news for many, there is a lack of knowledge about news sites in terms of interaction-experience constructs and their structural relationships. The project aimed to examine people's use of news sites from the perspective of interaction-experience research by developing a model and, based on this model, to provide guidance for designers of news sites. The project comprises three research phases: (1) exploratory phase, (2) modelling phase and (3) experimental phase. In the exploratory phase, a review of literature and an exploratory study of interaction experience with news Web sites were conducted. The latter explored how users of a particular news site interact with the site and which aspects of their experience they report. Data for the exploratory study were collected with an online questionnaire and by recording participants' use of a news site under think-aloud instructions. In the modelling phase, an online questionnaire was used to collect answers to psychometric scales that were selected based on the literature review and the exploratory study. A measurement model was formulated to test the relationship between measurement items and the measurement scales, and structural models were formulated to test hypotheses related to the structural relationships of variables. Following the test results, a model of interaction experience with news sites was formulated to predict outcome measures of interaction experience from variables measuring aspects of interaction experience. Components of interaction experience, in turn, were predicted from measures of perceived news-site characteristics. In the

experimental phase, an experiment was conducted to test the model of interaction experience with news sites in a controlled setting. Additionally, measures of person- and context characteristics were included in the prediction of components of interaction experience. The model of interaction experience with news sites was supported and accounted for a medium to substantial amount of variance in outcome measures. Finally, design guidance was derived from the model to advance interaction-experience knowledge, and conclusions were drawn regarding the model, in relation to existing research.

DEDICATION

I dedicate this thesis to my beloved mother, Aranyi Attiláné, and to the loving memory of my father, Aranyi Attila.

ACKNOWLEDGEMENTS

These past three years had been a very busy period in my life. Getting this far with my studies was not easy, and I have definitely not done it all by myself. First and foremost, I am deeply indebted to my director of studies, Paul van Schaik, for his constant and invaluable guidance and support throughout the whole process. I am very grateful to my second and third supervisors, Philip Barker and Chris Colbourn, for their help and their many useful comments on this manuscript, and to Andy Price for advice and guidance. I learned a lot from you all, and still have a lot to learn.

I couldn't even have started my studies without the Research Degrees Scholarship offered by Teesside University, let alone getting to submit this thesis. I would like to thank people at the Graduate Research School of Teesside University, especially Martin Leyland and Andrew Rawnsley, for organising research- and training events, and their tremendous help in administrative matters. Many thanks to Julie Martin at the Evening Gazette, and to many others at various departments of Teesside University and other universities who helped with recruitment to my studies. I would also like to thank Sziky Viola for her help in the graphics presentation of a poster. Thanks to all the participants who participated in my studies for their time and effort to provide valuable data.

Many thanks to fellow PhD students, especially Joseph Onibokun and Raza Habib, for all the help and the lively, often hilarious discussions. It was great fun working with you. Also thanks to Georgios Antonopoulos and Georgios Papanicolaou, both very knowledgeable researchers and extraordinarily kind persons, for the many refreshing coffee and discussions. I would like to extend my gratitude to Joseph Kean and Tom Gee, my fellow powerlifters, who helped me become a much better

sportsman. Training together helped me turn my mind away from my research when I needed rest, and attending (and winning) competitions were a significant part of my ‘PhD experience’. Last, but definitely not the least, thanks to my family for all the support and sacrifice, and to the many friends for being there and making these years a more pleasurable part of my life. I deeply appreciate all the love and help all of you gave me.

Thank you

ABBREVIATIONS

AC	Accessibility
AI	Adequacy of information
ANCOVA	Analysis of covariance
AVE	Average variance extracted
B2C	Business-to-client
BI	Behavioural intention
C2C	Client-to-client
BBC	British Broadcasting Corporation
CA	Classical aesthetics
CAN	Computer anxiety
CP	Computer playfulness
CPU	Central Processing Unit
CR	Composite reliability
CUE	Components of user-experience
CVPA	Centrality of visual product aesthetics
EA	Expressive aesthetics
HQ	Hedonic quality
KMO	Kaiser-Meyer-Olkin measure of sampling adequacy
M	Mean
ME	Mediated effect; proportion of the total effect that is mediated
n	Number of cases
N	Total number of cases
NEG	Negative affect
NHS	National Health Service

ns	Not statistically significant
NZH	New Zealand Herald
ONS	Online news service
OS	Operating system
PANAS	Positive and negative affect schedule
PAT	Person-artefact-task
PCSQ	Perceived core service quality
PD	Perceived disorientation
PE	Perceived enjoyment
POS	Positive affect
PSSQ	Perceived supplementary service quality
PLS	Partial least squares
PQ	Pragmatic quality
PUID	Perceived user-interface design
SD	Standard deviation
SE	Standard error
sr^2	Semipartial correlation squared
SUMI	Subjective usability measurement inventory
TAM	Technology acceptance model
TPB	Theory of planned behaviour
UC	Usefulness of content
UTAUT	Unified theory of use and acceptance of technology
UX	User-experience
VIF	Variance inflation factor

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

Chapter 1

Introduction

Contents

Chapter 1: introduction

1.1. Overview

The primary aim of this chapter is to present a background and rationale to the research undertaken and presented in this thesis. The chapter starts with a brief discussion of the increasing role and presence of experiential factors in research related to interaction between humans and interactive technologies. Next, news Web sites, as a specific type of interactive system, are introduced and a need for research is identified. Four broad research questions are formulated and the aims of the current research are designated. The chapter closes with an outline of the structure of this thesis.

1.2. Experiential factors in human-computer interaction

Although the discipline of human-computer interaction (HCI) was originally aimed at the study and design of interaction between people and computers (Card et al., 1983), research in this field in the past three decades encompassed a broad range of interactive devices and technologies, such as automated teller machines (Haynes & Thies, 1991; Tractinsky et al., 2000), portable electronic devices (e.g., personal digital assistants, portable audio players and mobile phones) and electronic performance support systems (Barker & van Schaik, 2010). A common property of these technologies that they are complex systems that afford for a wide variety of activities; therefore, the application of these systems is more accurately described as ‘interaction with an artefact’ rather than ‘use of a tool’.

Research in HCI traditionally focused on the achievement of behavioural goals in work-related settings (Hassenzahl & Tractinsky, 2006). Influential empirical and theoretical work in the late 1980s regarded the acceptance and use of information systems and computer applications (e.g., electronic-mail clients and word

processors) (Davis, 1989). A main issue with the application of information technology in the early years of human-computer interaction was the reluctance of users to invest effort to learn and use available systems that held the promise to substantially increase their performance at the workplace (Carroll & Rosson, 1987). Therefore, the usefulness and usability of computer systems were in the fore of human-computer interaction research, system design and evaluation. However, with the rapidly increasing proliferation of personal computing since the 1990s, information technology spread outside the workplace-context and was adopted for alternative uses, such as leisure (e.g., computer-gaming and instant messaging; van der Heijden, 2004), Internet use (Lederer et al., 2000), electronic retail and marketing (Barwise et al., 2002) and media consumption (e.g., online news; Chen & Corkindale, 2008). Consequently, experiential factors, such as aesthetics and enjoyment, began to receive increasing attention in human-computer interaction (Alben, 1996; Hassenzahl & Tractinsky, 2006).

'User-experience' (frequently abbreviated as UX), or interaction experience, has become a major area of research in human-computer interaction. Interaction experience stresses that interactive products and systems do not only deliver functional benefits, they also deliver experiences, and users' intention to (re)live positive experiences is an important drive of technology use (Hassenzahl, 2003). The main aim of interaction-experience research is to describe how positive experiences with interactive technologies can be promoted (Law & van Schaik, 2010). Utilitarian aspects, such as the ease of use and learnability of a particular interactive system, are widely regarded as important, but insufficient by themselves to give a complete account for the acceptance, use and successfulness of interactive technologies (Dillon, 2001; Thüring & Mahlke, 2007). Experiential and non-utilitarian

aspects in interaction-experience research include, for example, mental workload (Zijlstra, 1993), flow (a state of total involvement) (Davis & Wiedenbeck, 2001), perceived enjoyment (Sun & Zhang, 2008), perceived aesthetics (Lavie & Tractinsky, 2004), affect and emotion (Mahlke & Minge 2008; Sun & Zhang 2008) and the overall judgement of an artefact's beauty (Hassenzahl, 2004).

There is empirical support for the idea that experiential factors are important predictors of acceptance, use and successfulness of interactive products. For example, Hartmann et al. (2008) found that increased engagement and pleasure experienced during interactions with Web sites can be more important predictors of Web-site preference than usability, depending on the context of use and the target population of users. Models of interaction experience describe the important determinants of positive experiences with a wide range of interactive devices and technologies, from can openers to mobile-telephone menu layouts, and in several instances Web sites (e.g., van Schaik & Ling, 2011; Zhou & Fu, 2007). However, there is no unified definition of interaction experience (Law et al., 2009; Roto et al., 2011). An important rationale behind research on interaction experience is the notion that the success of interactive technologies is fundamentally connected to their ability to promote high-quality experiences. Consequently, a deeper understanding of the determinants and their relationships of positive experiences with interactive technologies is expected to guide successful system design.

1.3. News Web sites

News sites are a specific type of interactive system. News media have been publishing on the Web increasingly since the second half of the 1990s, and since then, Internet portals have become a favoured source of news for many (Allan,

2006). According to Hall (2001), the news sector was the third global professional sector to go online, after the military and academic sectors, and most major news providers in the Western world launched online versions of their newspapers within the first 18 months from the start of online news publishing in the middle of the 1990s. Fifteen years later, most news sites are no longer merely online versions of print newspapers. The Internet is not just another medium for journalism; online news has a great influence on how news stories are created and published, on the roles of journalists and readers (e.g., collaborative publishing; Bruns, 2005) and how people access information. A great and increasing number of news sites and online magazines are available on the Web with a wide variety of scope (e.g., international, national and local) and thematic focus (e.g., general, finance and entertainment). For example, as of December 2011, onlinenewspapers.com (2011) listed 578 news sites for England alone (other parts and territories of the United Kingdom excluded). Online news has an increasing penetration in the news sector and plays a significant role in how people acquire news and information about the world (Nguyen, 2008). Developers expend a growing amount of effort in enhancing the interactivity of news sites, thereby promoting user-generated content and communication among users, in accordance with a trend described as Web 2.0 (O'Reilly, 2007). Because of the immediacy of online news, that is, the short time it takes for news stories to get from news producers (and content-generating users) to audiences, print newspapers and television are replaced as the primary sources of breaking news (Hall, 2001).

According to Foust (2009), advantages of online journalism over traditional print media include the following. First, readers of news sites have a greater degree of freedom of choice as to which pieces of information they retrieve and when, due to the practically constant availability of online information and access to them through

(portable) Internet-enabled devices (e.g., mobile phones and tablet PCs). Second, readers can access information in a non-linear manner, due to the hyperlinked structure of news sites. Third, the amount of information accessible on news sites is not limited by the physical size of printed newspapers and even old news entries can be readily available. Therefore, news sites can successfully overcome the spatial and temporal limitations of traditional print media. Fourth, news sites allow for information to be published and updated almost instantaneously (i.e., continuously or according to a '24/7' news cycle), and enable journalists to include multimedia content (e.g., sound and video) in news stories. Fifth, news sites provide the potential for interactivity and user-generated content, such as discussion forums and the users' ability to comment news entries. As a consequence, the formerly passive readers can take part in the production of news (i.e., participatory journalism; Deuze et al., 2007).

The majority of online-news providers follow the business model of providing free access to news for users and generating revenue by selling advertisement space[1]. Therefore, users can access news on the Internet quickly, practically anytime and anywhere and, in most cases, essentially free of charge. With the world-wide trend of increasing use and penetration of the Internet, and the spread of handheld devices with constant access to the Web through wireless connection, such as netbooks, smart phones and tablet computers, ease of access to news and information in general is increasing. For example, in the United Kingdom, as of 31 March 2011,

[1] However, there is a tendency in the online news sector to move towards a subscription-based business model. A notable example is the News Corporation media conglomerate (News Corporation, 2009), which switched to subscription-based news distribution for a number of products, for example, The Times and The Sunday Times in the United Kingdom, in June 2010.

82%[2] of the population was estimated to have access to the Internet (Internet World Stats, 2011). According to Rosenstiel and Mitchell (2011), 47% of the adult population in the United States reported to access at least some local news and information on their mobile phone or tablet computer. According to a survey based on a nationally representative sample of households by the Office for National Statistics (2011), 77% of households in Great Britain had access to the Internet and 45% of users used a mobile phone to access the Internet.

Although a majority of the most popular news sites are associated with a traditionally print or broadcast news organisation (e.g., BBC and The Guardian), there are also other types of online news service (Foust, 2009). For example, news-aggregator sites compile news stories and links tailored to users' interest, specifications and/or browsing information rather than creating and publishing genuine content (e.g., Google News and My Yahoo!). Instead of as content generators, aggregators can be considered as indexes to a large pool of online news by providing links, headers and editorial overviews to other sites that actually produce news content, although some news providers are content generators and aggregators at the same time (Hall, 2001). Another example is hyperlocal online news, which provides content for geographic areas that are not covered well by major news providers. As a consequence, the content provided by hyperlocal sites is typically relevant and specialised to a limited group of readers, such as the residents of a town or community. For the purpose of the current thesis, only sites that undertake content generation are considered as news sites, regardless of scope (e.g., national and hyperlocal) and thematic focus. The main reason for excluding aggregators is that

[2] Note that estimates of different sources may vary, based on sampling and data-collection methods.

users following links to news entries on aggregator sites generally arrive at other news sites, which presents methodological problems in studying users' interaction experience with particular news sites.

1.4. Research questions

Hall (2001) argues that the monopoly of large media conglomerates (such as General Electric, The Walt Disney Company and News Corporation) is harder to defend on the Internet than on the largely monopolised print and broadcast media, due to large potential number of sources of information and the relative ease and low cost for anybody to publish on the Web; therefore, these companies are forced back into a more competitive business environment. As a consequence, transparency (Karlsson, 2011), trust in news providers (Kohring & Matthes, 2007) and experiential factors (e.g., enjoyment and mental workload) may pay a greater role in the successfulness of news Web sites. Therefore, news sites that provide high-quality experiences are expected to promote on-going, repeated interaction. However, despite the growing importance of online news publishing in the media sector and a shifting focus towards experiential factors in human-computer interaction, there is a lack of academic research on news sites in terms of interaction-experience constructs and their structural relationships. Although the news sector has had an increasing Web presence in recent years, there is a lack of knowledge about how news sites can be designed to promote a high-quality interaction experience. In particular, there is a lack of knowledge as to which characteristics of news sites, which can be seen as being used for both information access and leisure (Flavián & Guerra, 2006), are important in the formation of users' experience with this type of interactive technology. Therefore, the purpose of the research project presented in this thesis is to examine people's use of news sites from a human-computer

interaction perspective, and to formulate and test a psychological model of interaction experience with news sites. A deeper understanding of the relationships between the experiential aspects of interaction, and certain characteristics of news sites, their users and the contexts in which interactions take place, is expected to inform system design. The project addresses the following research questions.

1. What are the factors of end-users' interaction experience with news sites?

2. How can the factors of interaction experience be measured?

3. How are the different factors of interaction experience related?

4. How can the knowledge acquired by answering Research Questions 1, 2 and 3 be applied to guide system design?

Research Question 1 can be addressed by examining available research literature and conducting user-tests with users of news sites. Once relevant aspect of interaction experience with news sites are identified, measurement instruments, such as validated psychometric scales, can be adopted from reports of existing research and new instruments can be developed if necessary. Research Question 2 can be addressed by formulating a measurement model of interaction experience which specifies the relationship between the observed indicators and the components of interaction experience (Pedhazur & Pedhazur-Schmelkin, 1991). The reliability, validity and factor structure of indicators of interaction experience are assessed in empirical research. Research Question 3 can be addressed by formulating a

structural model which specifies the relationship between the components of interaction experience (Pedhazur & Pedhazur-Schmelkin, 1991). The formulation of the structural model is based on tests of hypotheses derived from previous research and theoretical considerations. Experimental testing of the structural model can provide further support for its comprehensiveness and predictive value. Research Question 4 relates to the application of the results of the research and the structural model to derive design guidance.

In summary, the expected impact of the research project presented in this thesis is twofold. First, the project contributes to human-computer interaction knowledge by the application of existing models to the domain of online news, and by formulating and testing a comprehensive model of end users' interaction experience with news sites. Second, by establishing and quantifying connections between designable product characteristics and measures of interaction experience, design guidance can be derived to aid developers of news Web sites.

1.5. The structure of this thesis

The four research questions presented in the previous section are addressed in six chapters in this thesis. Research Question 1 is addressed in Chapters 2 and 3. Chapter 2 outlines the concept of interaction experience for the purpose of this thesis, presents a literature review of models of technology acceptance and interaction experience in human-computer interaction research and identifies aspects of interaction experience of potential relevance to the current research. Chapter 3 presents an exploratory study that identifies self-reported aspects of interaction experience with news sites. Research Question 2 is addressed in Chapter 4, which presents a selection of variables to be included in a measurement model, data

collection for these measures and the measurement model itself. Research Question 3 is addressed in Chapters 5 and 6. Chapter 5 presents a structural model of interaction experience with news sites and hypothesis tests related to it, while Chapter 6 presents an experimental test of the structural model. Research Question 4 is addressed in Chapter 7, which provides a summary of the research presented in this thesis and derives design guidance from the studies.

Three separate studies are presented in this thesis. Study 1, presented in Chapter 3, applied an online questionnaire and a think-aloud protocol to explore participants' use of a particular news site. The results of Study 1 were used to guide the selection of measures in Study 2. Study 2, presented in Chapter 4, applied an online questionnaire to collect responses from a large number of users of a range of news sites. The data collected in Study 2 served as the basis of the measurement model presented in Chapter 4 and the structural model presented in Chapter 5. Study 3, presented in Chapter 6, applied an experimental setting in a controlled environment to complement and test the structural model. Finally, a discussion of the studies, their limitations and suggestions for future work are presented in Chapter 7.

Chapter 2

Psychological models of technology acceptance and interaction experience

Contents

2.1. Overview

This chapter presents a literature review that aims to provide a general framework for the thesis and a review of various psychological models that incorporate important variables of interaction experience. Note that the research in which these models have been proposed and investigated do not use the term 'interaction experience', and in most cases users' experience of interacting with a computer system is not addressed directly. The constructs of the models that are to be reviewed here, however, are considered relevant to interaction experience.

The chapter starts by defining interaction experience and usability for the purpose of this thesis, and specifying the use of these terms during further discussion. The principal components of human-technology interactions are identified, to be applied as a framework for examination and evaluation of the models presented in this chapter. Next, models of technology acceptance and adoption are introduced, in order to provide a historical and conceptual background for models of interaction experience and, more important, to present constructs and relations between them that are relevant to models that address the experience of users directly. Following this, models or interaction experience are presented. The chapter closes with the discussion and evaluation of the models of interaction experience and research related to them.

2.2. Interaction experience

In this thesis, interaction experience is used synonymously with user-experience or UX for short. The term *interaction experience* is used to describe the experience of users while interacting with interactive systems. The reason for preferring this term is that it focuses on users' experience in the context of interacting with technology;

therefore, it is more accurate than the term *UX*. The word *interaction* in 'interaction experience' implies interaction with a system, rather than the use of a tool. Furthermore, the term *user-experience* is fundamentally flawed, because research in this area typically investigates the experience of interaction (or use), not the 'experience of user'. Therefore, user-experience models are referred to as models of interaction experience in this thesis. The term *user-experience* or *UX* will be used in cases where reference is made to the name of existing models that use this term.

Although some authors give their explicit definitions of interaction experience (e.g., Hassenzahl & Tractinsky, 2006; Hassenzahl, 2008), a unified definition is lacking, despite attempts to collect existing definitions and viewpoints from experts and researchers of the field (Law et al., 2008). However, there seems to be a general agreement about the nature and scope of interaction experience (Law et al., 2009).

The fundamental, widely-accepted properties of interaction experience are dynamism, context-dependency and subjectivity (Law et al., 2009). *Dynamism* suggests that experience is not merely an end-product of interaction. Instead, it is an on-going phenomenon that changes over time. *Context-dependency* suggests that interaction experience is sensitive to changes in the setting in which it takes place. In other words, the quality of experiences is affected by contextual changes. Whether changes in the quality of experiences are contingent on certain contextual changes is subject to enquiry (e.g., mode of use and time pressure). *Subjectivity* refers to the notion that experience depends on the person who experiences as well as on contextual factors; therefore, it changes from person to person. Subjectivity does not mean, however, that the experiences of individuals cannot be compared, because everybody is different and, consequently, experiences everything uniquely (see also

Hassenzahl et al., 2010). Subjectivity merely emphasises the subjective side of product use and suggests that a certain variability of experiences is expected to be due to individual differences and preferences.

Hassenzahl's definition of interaction experience as "a momentary, primarily evaluative feeling (good-bad) while interacting with a product or service" (Hassenzahl, 2008, p. 12) is consistent with the three properties proposed by Law et al. (2009). The definition emphasises the subjective side of use, shifting attention from artefact characteristics to subjective feelings. *Dynamism* is expressed in the momentary nature of the evaluative feeling; interaction experience is seen as a present-oriented phenomenon that changes over time. Hassenzahl also notes that this present-oriented nature of interaction experience (and experiences in general) does not necessarily mean that it cannot be described in summarised retrospective accounts. It merely stresses that "the primary object of judgment remains the stream of passing momentary feelings" (Hassenzahl, 2008, p. 12). Although the context-dependent nature of interaction experience is not expressed in this definition, it is addressed and supported in experiments applying Hassenzahl's UX model (e.g., Hassenzahl & Ullrich, 2007).

Although interaction experience emphasises the momentary nature of experiences with technology, in other words, experiences are thought of as emerging during episodes of interaction with technologies, it involves a broader time-scale, which has important implications to measurement (Roto et al., 2011). Accounts of interaction experience can be elicited before use (anticipation of experience or presentation without interaction), during use (momentary account), directly after use (episodic account) and after an extended period (cumulative account). Furthermore, people

may have indirect experiences before interacting with a particular technology, based on, for example, previous experience with similar technologies, a presentation of the technology (e.g., advertisement and observation) or others' report or opinion.

Research suggests that on-line or real-time accounts of experience and retrospective evaluations may differ substantially. For example, Redelmeier and Kahneman (1996) found discrepancies between medical patients' real-time and retrospective accounts of pain. Furthermore, the authors found that the patients' retrospective accounts of pain were influenced predominantly by the peak intensity of pain during the medical procedure and by the intensity of pain at the end of the procedure (i.e., 'peak-end' rule), rather than the overall amount of pain experienced during the entire procedure. In other words, the retrospective account was not merely a sum or average of the real-time account.

Furthermore, research suggests that retrospective accounts or memories of experience may influence future decisions and behaviour to a greater extent than accounts of real-time experience. For example, applying long-term measurement in an ecological setting, Wirtz et al. (2003) compared anticipated, on-line and retrospective accounts of students' spring-break experience, and found that only retrospective accounts were directly predictive to the desire to repeat the experience. Although momentary accounts may be more accurate measures of actual (or 'objective') experience (Kahneman, 1999), retrospective measures (i.e., episodic and cumulative accounts) may be better predictors of future behaviour (Kahneman, 2000). In other words, applied in the present context, memories of experience may be more useful and accurate predictors of interaction outcomes, for example,

satisfaction, choice of alternative artefacts and future behaviour, as opposed to momentary or 'on-line' accounts of experience.

In addition to time-scale considerations in the measurement of experience, measures of interaction experience may involve a wide range of methods, for example, from the measurement of physiological responses as correlates of affective reactions (Mahlke & Minge, 2008) to approach-avoidance responses operationalized as attitudes toward a particular technology (Porat & Tractinsky, in press). As a consequence, studies related to, and models of, interaction experience may vary in terms of the time-scale and abstractness of measurement (see also Law & van Schaik, 2010). There is a wealth of standardised psychometric instruments available in reports of human-computer-interaction and interaction-experience research. Therefore, summarised retrospective accounts of interaction quality measured with the use of psychometric methods are proposed as a convenient way to measure interaction experience for the purpose of formulating and testing a psychological model of interaction experience in this thesis.

Frameworks for experience with artefacts have also been proposed outside human-computer interaction, for example, in the field of design (e.g., Norman, 2005; Desmet & Hekkert, 2007). Norman (2005) distinguishes visceral, behavioural and reflective levels of emotional design, while Desmet and Hekkert (2007) proposed aesthetic experience, emotional experience and experience of meaning as components of product-experience.[3] Hekkert (2006) defines product-experience as "the entire set of

[3] Desmet and Hekkert (2007) do not include usability among components of experience, because they consider usability as a source of experience which can generate and influence each of the three components.

effects that is elicited by the interaction between a user and a product, including the degree to which all our senses are gratified (aesthetic experience), the meanings we attach to the product (experience of meaning) and the feelings and emotions that are elicited (emotional experience (p. 160)".

For the purpose of this thesis, interaction experience within human-computer interaction is defined as *a user's evaluative reflection on his or her cognitive-emotional state while interacting with an interactive system or product, attributed to the use of the system or product*. This definition characterises interaction experience as subjective accounts of interaction quality that are situated in the process of interaction, emphasising the present-oriented nature of experiences. The attribution part of the definition indicates that the evaluative reflection on one's cognitive-emotional states is related to the interaction with the system or product. In other words, evaluative reflections are derived from the use of product or service, as opposed to circumstances that might affect the internal states of the users (for example, distracting background noise during interaction). Although the momentary (formative) nature of interaction experience is acknowledged in this definition, prospective (anticipated) and retrospective (summative) accounts of experience are also of theoretical and practical relevance, in line with the argument regarding methodology and prediction-relevance presented previously in this section.

Cognitive-emotional states can be characterised with subjective feelings during the interaction, such as pleasure, satisfaction and engagement. There is an important distinction to be drawn between pleasure and satisfaction: *satisfaction* involves an expectation about the interaction (Hassenzahl, 2003; Demir et al., 2009). For example, if a particular user expects to find information about a particular topic of

his/her interest on a news site, logs on to the site and finds a detailed discussion of this topic, it is expected that he/she will be satisfied. In the opposite case, if he/she finds no relevant information regarding his/her interest, it is likely to result in dissatisfaction. However, he/she may experience *pleasure* despite being dissatisfied about the unfulfilled expectation, because of, for example, finding something even more interesting on the news site that fulfils his/her need for stimulation. Pleasure may also be derived from the intrinsic pleasantness of an artefact (Demir et al., 2009). For example, the impressive graphics design of a news site may be pleasing to the senses, resulting in a certain level of pleasure during the interaction, without the possible fulfilment of needs or goals. This intrinsic pleasantness of the design corresponds to Norman's (2005) visceral level of emotional experience, whereas the expectation of behavioural or experiential outcomes and the subsequent satisfaction or dissatisfaction corresponds to Norman's behavioural level of emotional experiences. Although pleasure and satisfaction are similar emotional responses, the main difference is that satisfaction involves a component of expectation, and the resulting subjective feeling is dependent on the fulfilment of this expectation. Finally, *engagement* addresses the level of involvement during the interaction; however, pleasure is also associated with engagement (Hartmann et al., 2008). Pleasurable, satisfying and engaging experiences with an artefact are expected to lead to a positive overall evaluation of the artefact and intention to use.

Although experiences vary from person to person, they can be classified according to underlying human needs (e.g., autonomy, competence and relatedness) that are fulfilled during the experience (Hassenzahl et al., 2010). For example, news-site readers may find certain topics interesting to read about. From time to time, they visit particular news sites to find out if there is anything new on their topic of interest, or

generally, to find out if there is anything interesting at all. In other words, readers visit news sites to find interesting information, that is, to get stimulated by the sites. If their need for stimulation is fulfilled, it may contribute to positive interaction experience. In contrast, if readers find no information of interest at all, they may get bored (i.e., understimulated) and leave the site. From a design perspective, linking experiences to underlying human needs suggests that identifying the relevant needs in the case of a particular interactive technology helps to understand where positive experiences come from. Therefore, linking experiences to needs allows for the promotion of positive experiences by addressing underlying human needs in the design process. Hassenzahl (2008) explicitly links the quality of user-experience to the fulfilment of human needs. In the second part of his definition, he states that "Good UX is the consequence of fulfilling the human needs for autonomy, competency, stimulation (self-oriented), relatedness, and popularity (others-oriented) through interacting with the product or service (i.e., hedonic quality). Pragmatic quality facilitates the potential fulfilment of be-goals[4]" (Hassenzahl, 2008, p.12). Hassenzahl et al. (2010) found a clear relationship between the fulfilment of universal needs and positive affect, and a strong link between need fulfilment and hedonic product-quality perceptions. The authors suggest that pragmatic quality (i.e., user-perceived usability) of interactive systems or products, as a 'hygiene factor', may strongly contribute to negative experiences if it does not reach a satisfactory level expected by users. Therefore, it is plausible to explore need fulfilment and usability as sources of interaction experience with news sites.

[4]Be-goals are derived from the fulfilment of human needs (e.g., relatedness), whereas do-goals are related to achieving certain behavioural goals (e.g., making a telephone call). According to Carver and Scheier (1989), do-goals are derived from be-goals and do-goals are instrumental in achieving be-goals.

2.3. Usability

It is important to clarify the term *usability* for the purpose of the current thesis, because it is an important factor in human-computer interaction; therefore, it appears frequently in research on interaction experience. Furthermore, the usability of interactive systems is expected to influence the quality of interaction with these systems. For example, van Schaik and Ling (2011) found a positive effect of usability on interaction experience in a study using a Web-based encyclopaedia, which, similar to news sites, is a typical example of information-presenting Web sites. Shackel (1991, p. 24) proposed the following definition for the usability of a system:

> "the capability in human functional terms to be used easily and effectively by the specified range of users, given specific training and user support, to fulfil the specified range of tasks, within the specified range of environmental scenarios."

In the technology-acceptance literature, usability is described by the concepts of *perceived usefulness* and *perceived ease of use. Perceived usefulness* is defined as "the degree to which a person believes that using a particular system would enhance his or her job performance" (Davis, 1989, p. 320). Chen and Corkindale (2008) suggested that the definition for *perceived usefulness* should be different for online news services from the one applied to conventional offline information systems. Their modified definition for *perceived usefulness* is "the degree to which a person believes that using the ONS [online news service] will most enhance his or her task outcomes" (p. 299). *Perceived ease of use* is defined as "the degree to which a person believes that using a particular system would be free of effort" (Davis, 1989,

p. 320). Furthermore, dimensions of user-perceived usability include, for example, the learnability and controllability of systems (Kirakowski & Corbett, 1993).

Another widely accepted definition of usability is provided by ISO 9241 Part 11 (ISO, 1998): "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (p. 3). This definition is differs from Shackel's in two main respects. First, it does not contain ease of use explicitly. Second, it also incorporates satisfaction into usability. As discussed in the previous section, users' satisfaction with a particular product is derived from a wide range of aspects of interaction, and usability is but one of them. For this reason, the term *usability* is used here according to Shackel's definition, with an emphasis on ease of use, because ease of navigation, or browsing, is a major component of the usability of Web sites (Nielsen, 2000). Browsing can be defined as "[…] an activity in which one gathers information while scanning an information space without an explicit objective" (Toms, 2000, p. 424). When browsing news sites, readers may lack specific information-goals and navigate through hyperlinks to 'forage' for articles and pieces of information of potential interest (Pirolli & Card, 1999), as opposed to goal-directed search behaviour, which usually involves the use of keyword-based search facilities (Nah & Davis, 2002). Consequently, promoting ease of navigation is important in the design of news sites.

2.4. The main components of interaction experience

Successful design of interactive systems is dependent on taking into account important factors of interaction. Therefore, the important factors in the course of interaction need to be identified. Naturally, the interactive system under consideration is one of these factors, because there has to be an artefact to interact

with for interaction to occur. Similarly, there can be no interaction without an interacting agent. In human-computer interaction, the interacting agents are users. Successful design requires much attention to human factors, because interactive systems are intended to be used by human users. Isolating the design from human factors is likely to result in insufficient fit between an artefact and its users, resulting in poorer quality of interaction. Another factor that needs to be considered in the course of interaction with interactive systems is the tasks or activities performed using the system. Interactive systems are always designed with functionalities and facilities to allow for the manipulation of the system by the users. Without manipulation, there is no interaction, just passive observation. Functionality is always present, which means there are tasks and activities that can be performed using the system. Tasks and activities can vary broadly, for example, from the painstaking completion of an online tax form to the entertaining exploration of a friend's photo album on a social-networking site. Finally, interaction always takes place in a context. For example, one can check e-mails using a mobile phone while riding a crowded bus to work, or at home in front of a desktop computer in the weekend. The same person is likely to have different mind-sets and different preferences in different situations, which are likely to influence his/her approach to particular tasks and activities performed with particular artefacts.

According to Shackel (1991), usability depends upon the dynamic interplay of the following four components: user, task, tool and environment (see Figure 2.1). In Shackel's conceptual framework, usability depends on (a) the design of a tool (system) in relation to its users, tasks and environments, and (b) the success of support provided to the users (e.g., trainings, manuals and help facilities). The first part of this framework can be broadened and applied to interaction experience:

interaction experience depends on the design of a tool (interactive system) in relation to the users (human characteristics), tasks and environments (context). The second part is excluded from generalisation to interaction experience, because user-support is conceptually more closely related to the usability of a system as an antecedent of *perceived ease of use* (Cho et al., 2009), the latter of which is considered as one of many contributing factors to interaction experience.

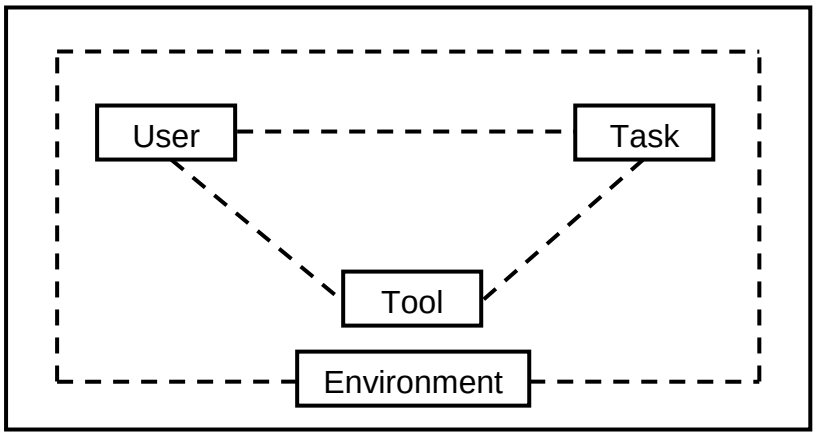


Figure 2.1. The four principal components in a human-machine system (Shackel, 1991).

Another model incorporating the factors of interaction experience identified in this section is the person-artefact-task (PAT) model, proposed by Finneran and Zhang (2003) (see Figure 2.2). In the PAT model, interaction experience is conceptualised using the concept of flow. Flow, or optimal experience, can be defined as a holistic sensation that people feel when they act with total involvement (Csikszentmihalyi, 1990). The PAT model conceptualises flow antecedents (the factors that lead to flow experience) in computer-mediated environments by addressing three distinct, but interacting components: person (P), artefact (A) and task (T). As a consequence of

an activity evoking flow, people may feel high levels of enjoyment and satisfaction, in other words, a high-quality interaction experience. Activity in the model involves task and artefact attributes, where task represents the main goal of the activity and the artefact represents the tool for accomplishing it.

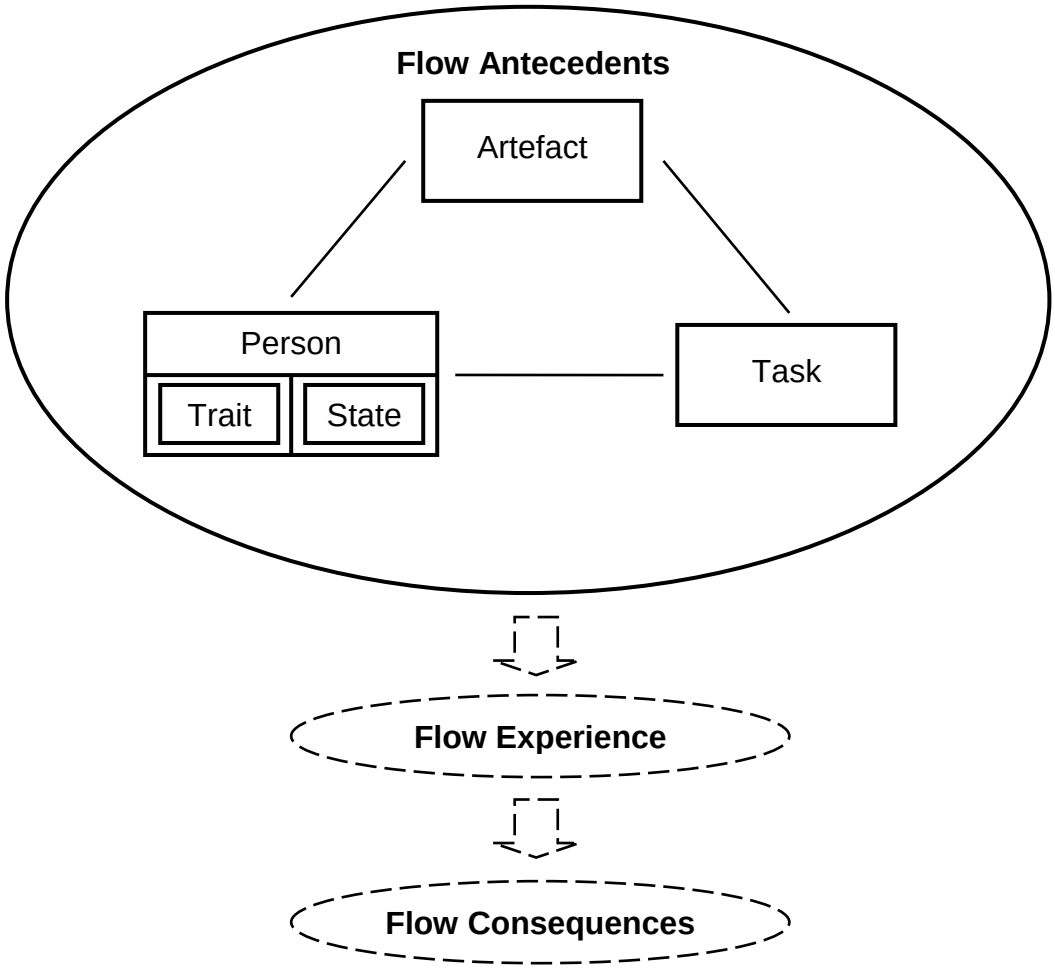


Figure 2.2. The PAT model (adopted from Finneran & Zhang, 2003)

Person characteristics in the model are broken down into state and trait attributes. Traits are considered to be effectively unchanging characteristics of a person (e.g., autotelic personality), whereas state attributes are dynamic and dependent on the particular environment (e.g., mood). The artefact factor in the model represents the properties of a particular interactive system with which a particular person interacts

while completing a particular task. The use of the term *artefact* is preferred, because it is broad enough to cover any activity that involves the use of a human-made system. It is a more neutral term than *tool* or *toy*, because the term *tool* implies a system used to achieve an external goal, while *toy* implies a system used for its own sake.

To what end a person uses an artefact is captured in the task factor in the PAT model. Interactive systems, such as news sites, can be used as a 'toy' or as a 'tool', in other words, for information access and leisure at the same time, depending on the way their users approach the interaction. This distinction is characterised in the concept of *mode of use* (Hassenzahl & Ullrich, 2007; van Schaik & Ling, 2009), which distinguishes two approaches a person can take while interacting with an artefact. In *goal mode*, the user is focused on achieving goals (e.g., searching for news to acquire information on a specific topic), while in *action mode* the user is focused on interacting with the product (e.g., browsing for news that may be interesting in general).

The PAT model also deals with the two-way interactions between the three factors (task-artefact, person-task and person-artefact) and three-way interactions. Task-artefact interaction, for example, may be described as task-technology fit (Dishaw & Strong, 1999), which implies the matching of the capabilities of the technology to the demands of the task, in other words, the ability of the artefact to perform a task. Thus, the artefact-person interaction in the PAT model is related to usability. Finally, person-task interaction encompasses elements from flow theory, for example, related to how clear a particular person perceives the task to be, how much control he/she experiences while completing the task and the amount of feedback he/she receives.

It is proposed here that the four principal components that have been introduced, along with the possible interactions between the elements, are general and broad enough categories to encompass every relevant aspect of interaction experience. Therefore, these main components will be used as a framework, in which the different models of interaction experience can be positioned, based on the components and interactions they represent in detail. The characteristics of a particular artefact, the characteristics of the task performed while interacting with the artefact and the characteristics of users, in turn, influence the process of interaction, which leads to consequences, such as the positive appraisal of the artefact.

2.5. Models of technology acceptance

2.5.1. Technology acceptance model

The technology acceptance model (TAM) is an adaptation of the theory of reasoned action (Fishbein & Ajzen, 1975) to information systems. It is specifically meant to explain and predict computer usage behaviour (Davis, 1986). During the past two decades it has been a highly useful model for understanding the use and adoption of new technologies, especially in work situations (Chen & Corkindale, 2008). It is considered a well-established, robust, powerful and parsimonious model for predicting user-acceptance (Venkatesh & Davis, 2000).

TAM posits that two main independent behavioural-belief constructs, *perceived usefulness* and *perceived ease of use*, determine an individual's use of a system, with *behavioural intention* as a mediator between the belief constructs and actual system use (Davis, 1989). The effects of external variables (e.g., system characteristics and training) on *behavioural intention* are mediated by *perceived usefulness* and *perceived ease of use*. *Perceived usefulness* and *perceived ease of*

use are distinct, but related constructs. *Perceived ease of use* has a direct influence on *perceived usefulness*, because a system is more useful if it is easier to use; the less effortful the system is to use, the more using it can increase performance. Originally, the technology acceptance model also included the attitude construct from the theory of reasoned action, but it has been removed later for parsimony (Venkatesh et al., 2003).

Venkatesh and Davis (2000) proposed an updated version of the technology acceptance model called TAM2 (see Figure 2.3). TAM2 model introduces two sets of constructs as determinants of *perceived usefulness* and *behavioural intention* and deals with how their influence changes with increasing experience over time with a particular system. These theoretical constructs include *social influence processes* (subjective norm, voluntariness and image) and *cognitive instrumental processes* (job relevance, output quality, result demonstrability and *perceived ease of use*). Encompassing these constructs, up to 60% of the variance in *perceived usefulness* was explained in the authors' longitudinal study applying computer systems in four distinct organisations.

Similar to *perceived usefulness*, *perceived ease of use* has also been detailed in further research. For example, Venkatesh (2000) introduced control (internal and external, conceptualised as *computer self-efficacy* and *facilitating conditions*), intrinsic motivation (conceptualised as *computer playfulness*; CP) and emotion (conceptualised as *computer anxiety*; CAN) as antecedents of *perceived ease of use*, and traced how their effects changed over time in a model based on the principle of anchoring and adjustment. According to this principle, anchors, or pre-existing

beliefs about using a particular system, serve as reference points that are adjusted with growing experience with the system.

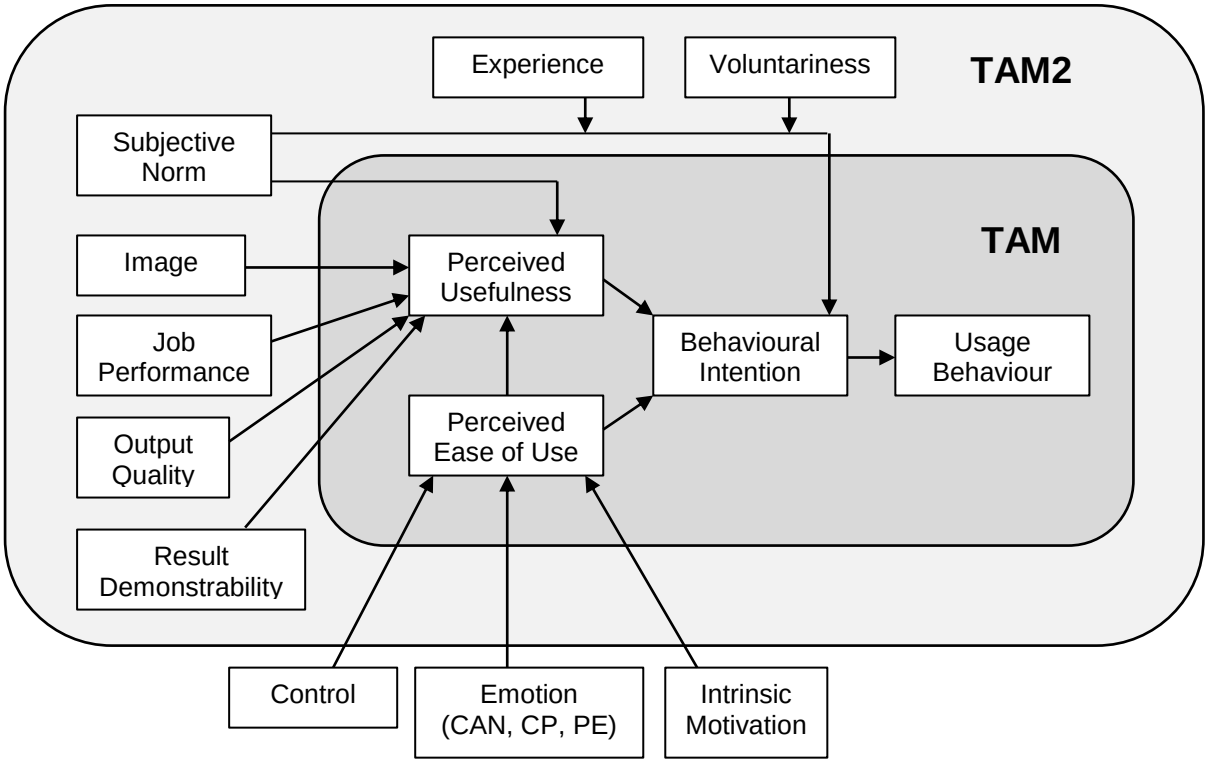


Figure 2.3. A technology acceptance model based on TAM and TAM2, augmented with additional variables.

Lederer et al. (2000) validated TAM for work-related tasks in the context of the World Wide Web and introduced a list of antecedents for *perceived usefulness* and *perceived ease of use*. Lederer and colleagues concluded that TAM can be considered a suitable theoretical foundation for the adaption of Web technologies. However, a number of Internet-related studies applying TAM concluded that *perceived usefulness* and *perceived ease of use* may not be significant predictors of intention to use in certain contexts (e.g., online retail), and other variables need to be introduced (e.g., enjoyment, trust, previous experience, consumer traits and situational factors) to better explain variability in *behavioural intention* and actual

system use (Monsuwé et al., 2004). For example, Gefen et al. (2003) integrated *trust* with TAM and identified antecedents of *trust* in the context of online retail.

TAM has undergone several extensions over the years. As Wixom and Todd (2005) point out, attempts to extend TAM have generally taken one of three approaches: (a) by introducing external variables as antecedents and moderators (e.g., demographics, system characteristics and personality traits) to the behavioural-belief constructs of *perceived usefulness* and *perceived ease of use*, (b) by introducing additional belief factors (e.g., *trialability* and *compatibility*) and (c) by introducing factors from related models (e.g., *subjective norm* and *perceived behavioural control*). TAM has also been merged with other models of information-technology utilisation. A notable example to this approach is the extension of TAM with constructs of the task-technology fit model (Dishaw & Strong, 1999).

Perceived usefulness and *perceived ease of use* have been complemented with a third belief construct, *perceived enjoyment* (PE), which is defined as "the extent to which the activity of using computers is perceived to be enjoyable in its own right, apart from any performance consequences that may be anticipated" (Davis et al., 1992, p. 1113). *Perceived enjoyment* is considered a state affect-variable; it is associated with the use of a particular technology at a particular time, thus it may change dramatically over time and across systems. Furthermore, in a technology acceptance model predicting the loyalty of users of mobile services, Cyr et al. (2006) found that design aesthetics was a significant predictor of *perceived enjoyment*, *perceived ease of use* and *perceived usefulness*. Sun and Zhang (2008) added *computer playfulness* as a trait affect-variable to an adaptation of TAM to the domain of information systems. As opposed to *perceived enjoyment* (as a state affect-

variable), *computer playfulness* (as a trait variable) is regarded stable over time. Sun and Zhang methodologically verified the mediating effects and causal relationships of these affect variables with *perceived usefulness*, *perceived ease of use* and *behavioural intention*.

Chen and Corkindale (2008) adopted the technology acceptance model to the application area of online news services. The core of the model is the original TAM, although perceived ease of use is replaced by *interface*, which is defined as "the extent to which a user finds the interface to be effective, fun and pleasant to use" (Chen & Corkindale, 2008, p. 304) (see Figure 2.4). Service quality, as an antecedent to *perceived usefulness*, is divided in the model into *perceived core service quality* (PCSQ), which encompasses the quality and presentation of the content, and *perceived supplementary service quality* (PSSQ); neither of these two are part of the core services, but both are closely connected to them (e.g., services associated with blogs, news alerts and search facilities). Subjective norm was reintroduced (it was excluded from TAM for parsimony) into the model as a social-influence variable, defined as a "person's perception that most people who are important to him think he should or should not perform the behaviour in question" (Fishbein & Ajzen, 1975, p. 302). Another important factor for adaption is *trust*, defined as "the extent to which a user believes the ONS [online news service] providers will fulfil their expectations and behave in a socially responsive manner" (Chen & Corkindale, 2008, p. 303). The model posits that reputable services invoke trust in consumers, enhancing usefulness perceptions and adaption. Interactivity of Web applications is addressed in the model by networking, defined as "the degree to which a user satisfies the online interactivities, including the B2C [business-to-client] and C2C [client-to-client], facilitated by an online service provider" (Chen &

Corkindale, 2008, p. 302). However, the adaptation of TAM to the application area of online news did not include psychometric measurement of the components of the model listed above; rather it was based on interviews with news media experts. Therefore, TAM was not validated in the context of online news.

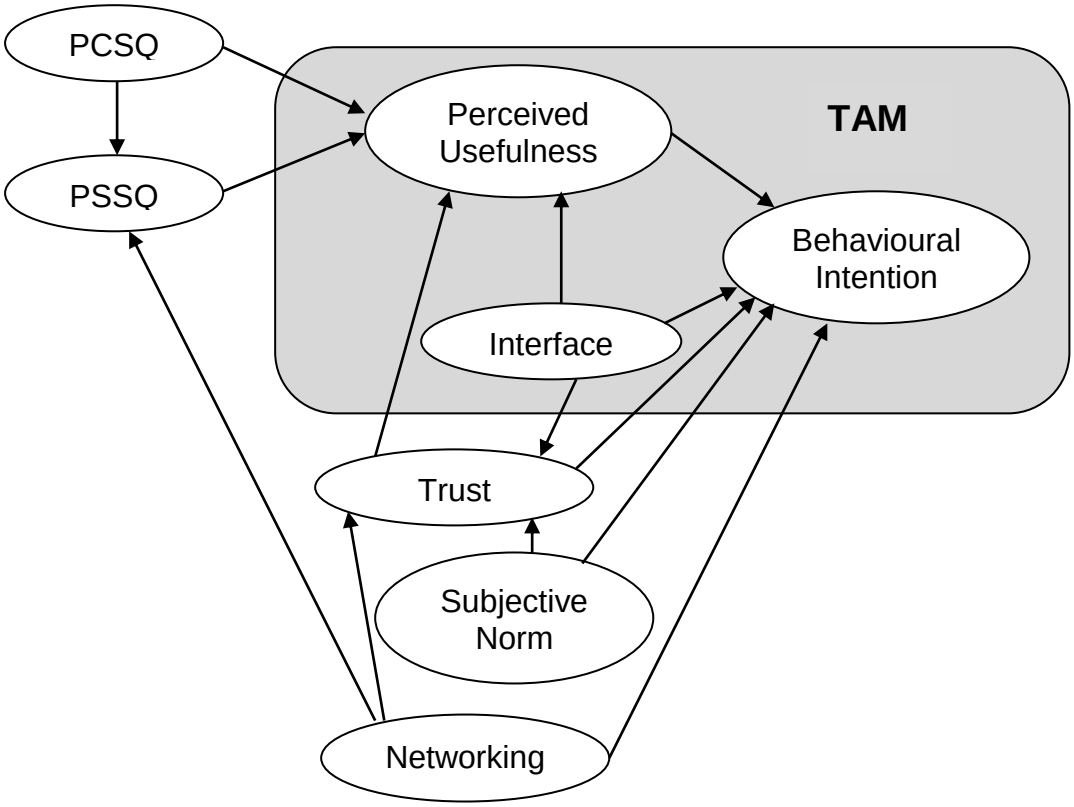


Figure 2.4. A model for adaption of online news services (adopted from Chen & Corkindale, 2008).

A recent adaptation of TAM to the field of Web technologies by Cho et al. (2009) emphasises the importance of user-interface design on continued usage intention (see Figure 2.5). User-interface design is defined as "the structural design of an interface that presents the features and instructional support of an information system" (Cho et al., 2009, p. 216). *Perceived user-interface design* is recognised as a major factor in Web-technology adoption, because the functionality of a computer system is reflected by the user-interface (Jaspers et al., 2004). The user-interface

acts as a point of contact between a particular user and the functions provided by the system. In other words, the interface is the 'face' of a particular system in the sense that it is the medium through which the system is presented to the user. With regards to interactive Web technologies, it is important that the interface can be adapted to individual needs and preferences. Adaptability can be described with the concept of *customisation*, which is defined as "the ability of users to modify the system according to their preferences" (Seneler et al., 2009, p. 495).

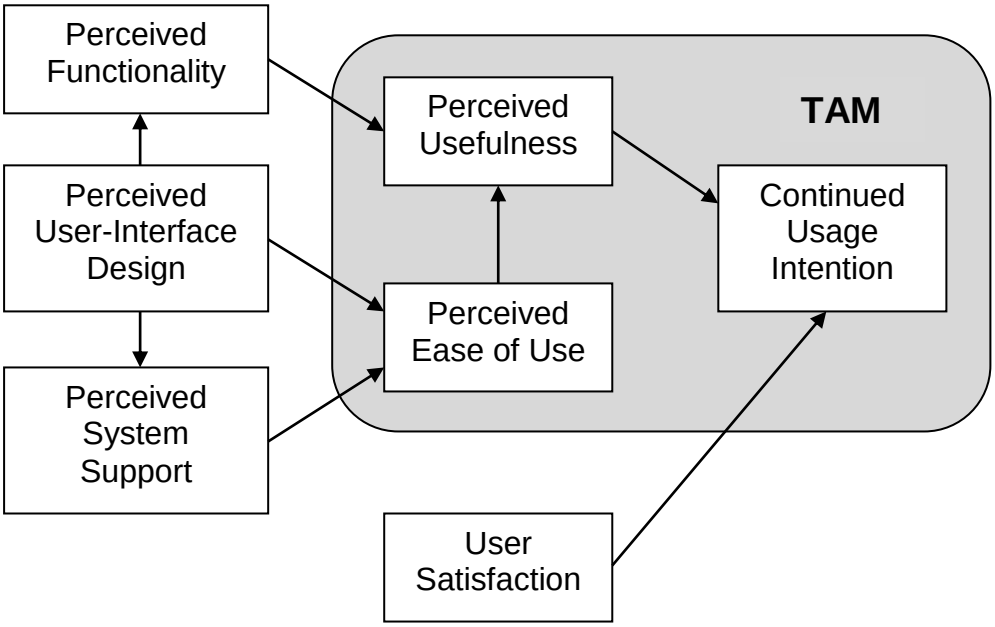


Figure 2.5. The role of perceived user interface design in continued usage intention (adopted from Cho et al, 2009).

In the model of Cho et al., the impact of *perceived user-interface design* on *perceived usefulness* and *perceived ease of use* is mediated through *perceived functionality* and *perceived system support*, respectively. *Perceived system support* is defined as "the perceived effectiveness of system support for [using] a system" (Cho et al., 2009, p. 219), including help desks, online support services and other facilities that enhance ease-of-use perceptions. The model also includes a *user-satisfaction*

variable with a direct influence on continued usage intention. *User-satisfaction* is defined as “an affective state representing an emotional reaction to the use of a technology” (Cho et al., 2009, p. 220). According to expectation-disconfirmation theory (Bhattacherjee, 2001), following one’s first-hand experience with a system, a satisfied user may have a high intention to continue using the system, due to the positive reinforcement of the attitude toward the system after using it (Cho et al., 2009). In other words, if after an initial trial, a particular user finds that the system was able to live up to his/her expectations and was of satisfactory help in attaining his/her goals, satisfaction with the system is expected to result in a positive attitude toward continued system use. Note that the effect of user’s satisfaction on *continued usage intention* is not mediated through the belief constructs of TAM. This variable is external to TAM in this regard and it can be viewed as a variable of experience that is not connected to usability, and it accounts for *continued usage intention* along with the usability-related variables of TAM.

In summary, TAM has become more and more detailed since its inception more than two decades ago and it has been adopted in several fields despite its origins in mandatory use of a technology within work situations. In the process of adoption, several new factors came into consideration, resulting in a growing list of antecedents to the original factors, belief constructs and various new determinants to intention to use. The original TAM is very parsimonious, but the experience of users eludes its scope, as it operates with belief constructs and, in some cases, the attitude construct, with all other effects mediated through them, and it deals with intention to use and system usage as outcome variables. Of course, it is a model of technology acceptance and not interaction experience. As presented above, many variables were introduced to uphold its explanatory power, reducing the simplicity of the

original model. Behavioural intention to using a particular system may not be explained fully through usability. This is explicitly represented in Cho et al.'s (2009) model where *user's satisfaction* is independent from *perceived usefulness* and *perceived ease of use*.

2.5.2. Unified theory of acceptance and use of technology

The unified theory of acceptance and use of technology (UTAUT) aims to explain the acceptance of technology and subsequent usage by incorporating different sets of technology-acceptance determinants from eight existing models of technology adoption into a unified model (Venkatesh et al., 2003). These models are the theory of reasoned action, TAM, motivational model, theory of planned behaviour (TPB), a combined TAM/TPB model, model of PC utilization, innovation diffusion theory and social cognitive theory. UTAUT was developed through a review and consolidation of the constructs of the eight models. These models were compared empirically in a longitudinal study. UTAUT was formulated on the basis of conceptual and empirical similarities between the models, tested with the data used in the assessments of the eight models and validated with new data sets. The validation of UTAUT found that the model accounted for 70% of the variance in intention to use.

UTAUT has four direct determinants of behavioural intention and actual usage behaviour: *performance expectancy*, *effort expectancy*, *social influence* and *facilitating conditions*. Four additional variables moderate the effects of these main determinants: *gender*, *age*, *experience* and *voluntariness of use* (see Figure 2.6).

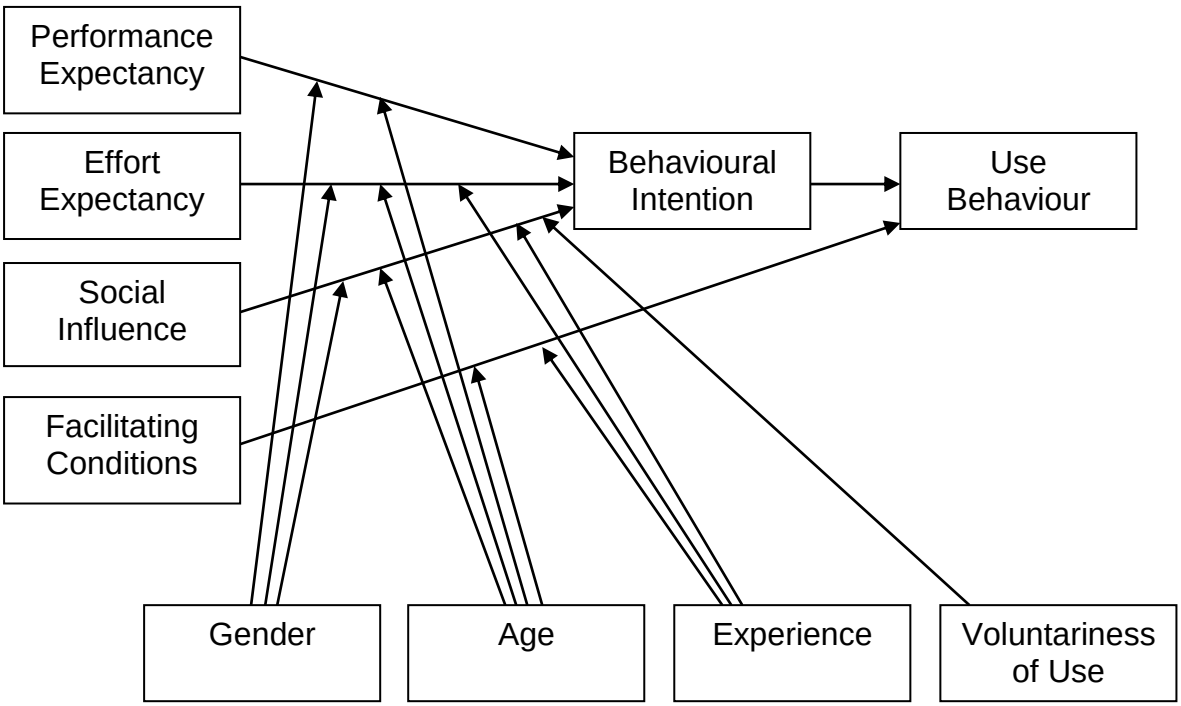


Figure 2.6. UTAUT (Venkatesh et al, 2003).

Performance expectancy is defined as "the degree to which an individual believes that using the system will help him or her to attain gains in job performance" (Venkatesh et al., 2003, p. 447) and corresponds to the *perceived usefulness* construct in TAM. *Performance expectancy* proved to be the strongest predictor of *behavioural intention* and remained significant with growing experience, both in voluntary and mandatory settings. *Effort expectancy* is defined as "the degree of ease associated with the use of the system" (Venkatesh et al., 2003, p. 450) and corresponds to the *perceived ease of use* construct in TAM. Effort expectancy had a significant effect on *behavioural intention* both in voluntary and mandatory settings, but it was found to lose this effect with experience.

Social influence is defined as "the degree to which an individual perceives that important others believe he or she should use the new system" (Venkatesh et al.,

2003, p. 451) and corresponds to *subjective norm* in TAM2. The effect of *social influence* on *behavioural intention* was found to be non-significant in voluntary settings. This effect was attributed to compliance in mandatory contexts, where individuals are expected to use the technology (Venkatesh & Davis, 2000). In mandatory settings, *social influence* became non-significant over time, due to the effects of internalization and identification of norms. *Facilitating conditions* are defined as "the degree to which an individual believes that an organizational and technical infrastructure exists to support use of the system" (Venkatesh et al., 2003, p. 453). *Facilitating conditions* was identified as an antecedent of *perceived ease of use* in TAM as part of the control factor (Venkatesh, 2000). *Facilitating conditions* proved to be a significant predictor of *use behaviour* both in voluntary and mandatory settings and its effect proved to become non-significant with experience. *Facilitating conditions* as a determinant of technology acceptance proved to have a direct effect on *use behaviour* beyond that explained by *behavioural intention* alone.

UTAUT provides a richer account of *behavioural intention* to use a system and *use behaviour* than the original TAM by incorporating a wide range of variables from other models. While the original TAM only included *perceived ease of use* and *perceived usefulness*, UTAUT incorporates these constructs as *effort expectancy* and *performance expectancy* and augments them with various person-related variables (gender, age and experience) and task and context-related variables (voluntariness of use, social influence and facilitating conditions) that are well-established in other models. The result is very similar to what TAM turned out to be after its gradual extension with additional variables. The effects of new variables were observed in the context of the original belief constructs to increase the variance explained by the model. Furthermore, the inclusion of variables such as *gender*, *age*,

experience and *voluntariness of use* as moderators allows for the examination of interactions between factors related to person, artefact and task/context characteristics in the model.

2.6. Models of interaction experience

2.6.1. User-experience model

Theoretical work related to the user-experience (UX) model emphasises the notion that interactive products are not merely means to end that provide functionalities in a usable manner in order to support effective and efficient task completion. Interactive products may also be regarded as a source of pleasure, joy, and enjoyment (Hassenzahl, 2003). Subjective aspects of product use, such as affective experiences during interaction with systems, are important determinants of choice and subsequent use of technologies. Interactive products not only deliver utilitarian benefits, they deliver experiences at the same time. Hassenzahl (2003, 2004 and 2008) argues that utilitarian concepts, such as usefulness and ease of use, are insufficient to account for the experience of users, although these factors are regarded as important facilitators of positive experiences.

According to Hassenzahl's user-experience (UX) model (see Figure 2.7), users perceive features of interactive products and construct their unique perceptions of product character. Product character comprises of pragmatic and hedonic attributes. Pragmatic attributes encompass utility and usability that allow for the manipulation of products. These attributes are also referred to as instrumental or utilitarian (Hassenzahl, 2006). Hedonic attributes are those factors that make interaction with a particular product pleasurable by fulfilling the human needs of autonomy,

competency, stimulation (self-oriented), relatedness and popularity (others-oriented) (Hassenzahl, 2008).

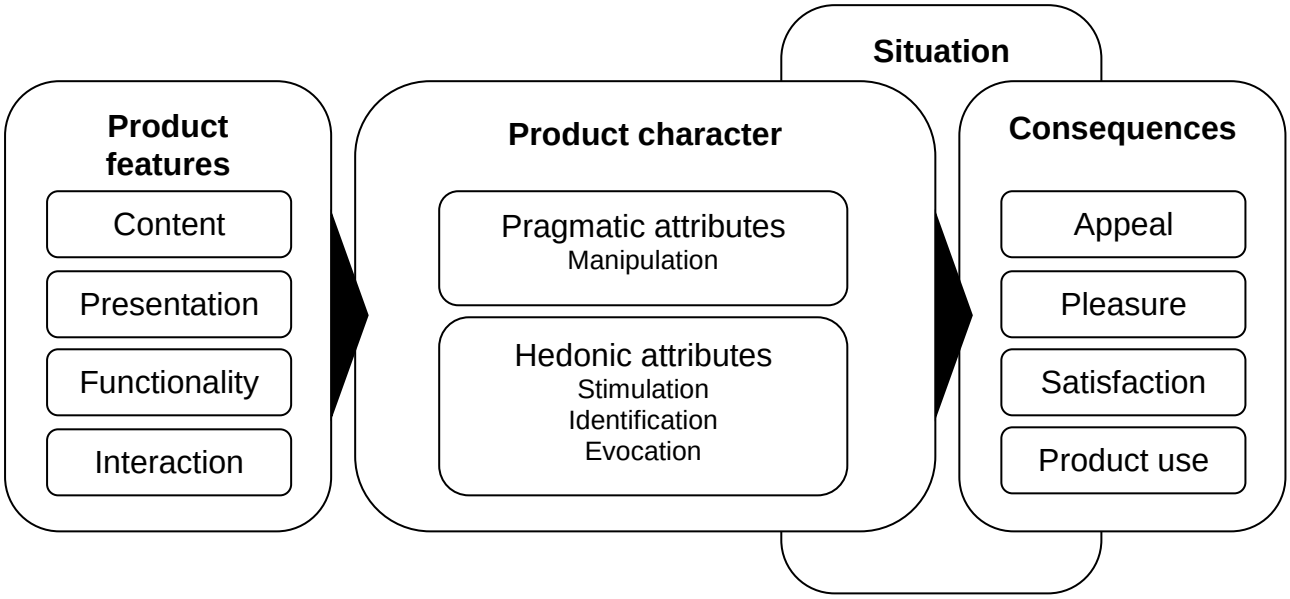


Figure 2.7. The user-experience model (based on Hassenzahl, 2003).

Hedonic attributes in the model are divided into three categories: providing stimulation, communicating identity and evoking valued memories. Pragmatic attributes emphasise the fulfilment of behavioural goals (do-goals), hedonic attributes emphasise psychological well-being (be goals) (Hassenzahl, 2003). Hedonic product quality attributes are related to be-goals, stand closer to the self and they are important drives for emotional product attachment (Hassenzahl & Roto, 2007). Perception of a particular product's character may lead to various consequences: judgement about the product's appeal (good/bad), emotional consequences (pleasure and satisfaction) and behavioural consequences (e.g., prolonged and subsequent use of the product). A product can be viewed as pragmatic, because it promotes effective and efficient manipulation of the environment, or it may be considered as hedonic, because it provides stimulation, identification or provokes

memories (Hassenzahl, 2003). Hassenzahl (2004) argues that because judgements of beauty can be derived from the appearance of products, they are more strongly related to beauty judgements after interaction than goodness ratings based on observation only to goodness rating after interaction. Goodness evaluations are more strongly influenced by pragmatic quality, and usability perceptions are strongly influenced by the actual use of interactive systems. The difference in the stability of goodness and beauty judgements before and after interaction is supported by the findings of van Schaik and Ling (2008), who tested and confirmed the user-experience model in the context of interaction with Web-sites. Furthermore, van Schaik and Ling (2011) tested an integrated model of interaction experience and technology acceptance using two versions of an information-presenting Web site varying in terms of usability and found that constructs of the user-experience and technology-acceptance models are separate underlying dimensions of users' experience. These findings emphasise the need to address a wide range of determinants in interaction-experience modelling to account for interaction outcomes, such as the successfulness of a particular interactive technology.

The consequences of perceived product characteristics are not always the same, but subject to contextual factors. In studies applying the user-experience model, context is conceptualised as mode of use (action mode or goal mode, as described in Section 2.4). In research, goal mode is characterised with the presence of externally given instrumental goals (i.e., tasks), whereas in action mode participants are instructed to use a particular system as they like (e.g., exploration and free browsing). Mode of use was found to affect interaction experience in several respects. For example, Hassenzahl et al. (2002) found that *pragmatic quality* (i.e., user-perceived usability) was more strongly correlated with the overall appeal of Web

sites in goal mode than in action mode. Hassenzahl and Roto (2007) found that the presence of active instrumental goals influence the experience of interaction and subsequent retrospective judgements. In an experiment using an interactive storytelling system (Hassenzahl & Ullrich, 2007), spontaneity experienced during the interaction was negatively related to affect and the appeal of the system in goal mode, whereas in action mode, spontaneity was positively related to affect and appeal. The interactive storytelling system in the study was evaluated in terms of usability in goal mode (i.e., the system's capability to support attainments of goals), but usability was not predictive to product appeal in action mode. These results suggest that participants changed their evaluation criteria according to the context of use. In two experiments using Web sites as artefacts, Van Schaik and Ling (2009), found evidence for the effect of mode of use in producing more stable judgements of aesthetics before interaction with Web sites, and concluded that context needs to be addressed in research on interaction experience. Overall, research applying the user-experience model demonstrated that perceptions of product attributes and evaluative judgements are context-dependent.

The user-experience model links hedonic attributes to the fulfilment of universal psychological needs. According to the model, hedonic attributes promote the fulfilment of be-goals and emphasise individuals' well-being (Hassenzahl, 2003). On the other hand, pragmatic attributes, such as usability, support the achievement of behavioural goals (do-goals). Hassenzahl (2008) identified the ability of interactive products to support need fulfilment as a source of affective reactions in relation to the use of interactive technologies. In a study based on the accounts of users of a wide range of interactive technologies, Hassenzahl et al. (2010) assessed the relationships between the fulfilment of universal human needs (e.g., competence,

stimulation, relatedness and popularity), positive affect and perceptions of product attributes. The study revealed significant relationships between need fulfilment and positive affect, and need fulfilment was strongly related to hedonic quality perceptions. Furthermore, in accordance with the distinction between hygiene factors and motivator factors (Zhang & von Dran, 2000), the role of usability as a hygiene factor[5] was supported. Studies applying the user-experience model suggest that experiences with interactive products can be categorised in terms of their ability to support need fulfilment.

The conceptual framework of the user-experience model outlined in Figure 2.7 has the merit of connecting the perception of product attributes to product features. From a designer's perspective, this allows for making connections between designable product characteristics and the intended product character, in order to control user-experience (at least to some extent) through intentional design choices. However, research related to the model is more strongly concerned with context (mode of use) and need fulfilment as a source of positive experiences with interactive products, and aims to provide a conceptual framework for studying interaction experience rather than to serve as a basis for design guidance for developers of particular technologies. The UX model also links hedonic attributes to needs and values, suggesting that designing for the fulfilment of these needs will lead to more positive interaction experience, making a particular product more successful and more appreciated by its users. In contrast to the models of technology acceptance and adoption discussed before in this chapter, in the user-experience model usability is

[5] According to Zhang and von Dran, the absence of hygiene factors leads to dissatisfaction, but their presence does not lead to satisfaction. The presence of motivational factors, on the other hand, leads to satisfaction and promotes the quality of interaction experience.

not the centre of attention and the model gives more credit to non-utilitarian factors, such as design aesthetics and the role of interactive technologies in communicating about the self (identification). These factors are regarded as at least as important determinants of users' experience. Indeed, experience itself is in the centre of the model; variables corresponding to attitudes toward the system and behavioural intentions in previously discussed models are regarded as consequences of the experience of individual person-technology interactions.

2.6.2. An environmental-psychology model of interaction experience

The environmental-psychology model[6] of interaction experience (Porat & Tractinsky, in press) aims to describe how design characteristics of Internet shopping sites influence the emotional states of users and how these changes influence attitudes towards these sites. The model is based on an approach from environmental psychology (Mehrabian & Russell, 1974), which posits that changes in the environment induce certain emotional changes, which in turn influence behaviour (approach/avoidance responses).

Web site characteristics (or environmental stimuli in general) in the model are described in terms of aesthetics and usability (see Figure 2.8). Visual aesthetics of computer interfaces (e.g., Web sites) are important determinants of pleasure and satisfaction for users (Lavie & Tractinsky, 2004). Visual aesthetics is particularly important in the case of user-interface design of Web technologies, because other modalities are limited in their capabilities (e.g., auditory) or completely unavailable (e.g., olfactory and tactile). Lavie and Tractinsky derived theoretically and

[6] The model was not named by Porat and Tractinsky (in press). It is referred to as environmental-psychology model in further discussion.

established empirically that there are two dimensions of aesthetics: classical and expressive aesthetics (Lavie & Tractinsky, 2004). *Classical aesthetics* represent clear and orderly design, whereas *expressive aesthetics* represent creativity, originality and innovativeness.

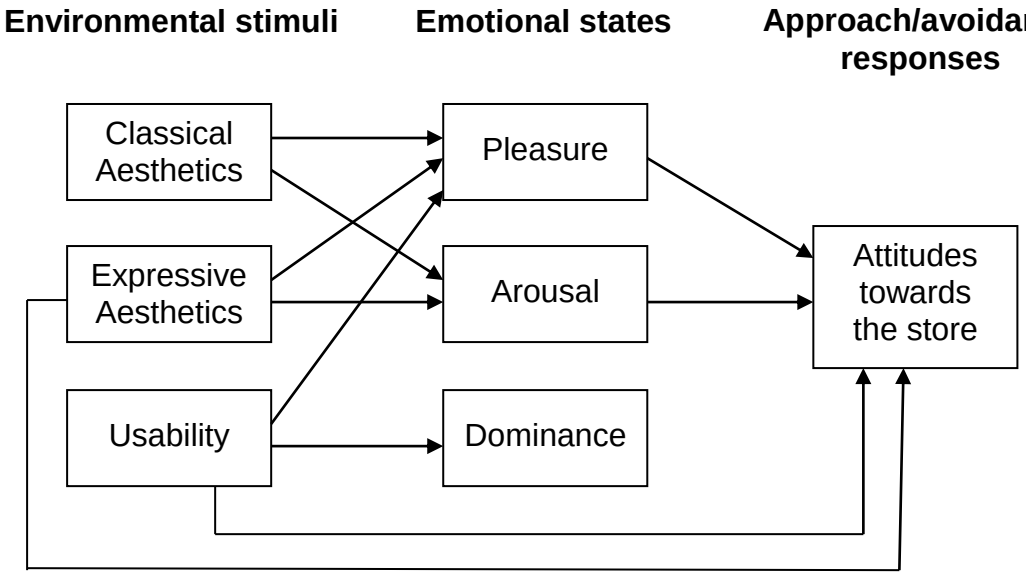


Figure 2.8. The environmental-psychology model (Porat & Tractinsky, in press).

Note. Arrows indicate significant paths in the model.

The model includes three dimensions of emotional responses (based on Mehrabian & Russell, 1974). *Pleasure* refers to the valence of the emotional response (positive or negative). *Arousal* refers to the degree of stimulation caused by the environmental stimuli. *Dominance* represents the degree of one's perceived control over the environment. According to the model, positive emotional responses (i.e., good interaction experience) are important determinants of positive attitudes towards Internet shopping sites (or artefacts in general). In the study that formed the basis of the model, aesthetics and usability were both important determinants of emotional responses. Usability had a significant positive effect on the dominance dimension of emotional responses, meaning that the higher level of usability resulted in better

perceived control in the course of interaction. Dominance, however, showed no significant effect on attitudes with the artefacts and research design used by the authors. The results suggested that increased usability made the interaction more pleasurable and made people feel like more in control. Being in control, however, may be something that is expected during the interaction with a Web store. It may be a minimum requirement, a hygiene factor (Zhang & von Dran, 2000), the fulfilment of which helps avoid dissatisfaction, but once fulfilled, it does not add more to users' interaction experience. On the other hand, usability had a significant direct effect on attitudes. Furthermore, the effect of expressive aesthetics on attitudes was partially mediated through the affect dimensions of valence and arousal, whereas the effect of classical aesthetics on attitudes was fully mediated through affect dimensions.

The findings of Porat and Tractinsky (in press) suggest that the strengths of the connections in the model are dependent on the type of stimulus or artefact under consideration. Expressive aesthetics may be more salient in the case of a certain product type, classical aesthetics and usability might be more salient for another. This notion may be generalised from Internet shopping sites to other online applications. For instance, the salience of classical aesthetics may be greater for a financial news site than for a 'soft-news' portal. Similarly, usability may be more important than aesthetics in the case of, for example, online document management applications.

In the model, interaction experience is characterised as emotional responses to perceived artefact characteristics (usability and aesthetics). The environmental-psychology model is not a general model of interaction experience, because it was explicitly formulated to account for approach/avoidance responses of users of online

retail sites. However, the model demonstrated that affective reactions were affected by aesthetic features and usability of retail Web sites, and affective reactions to perceived product characteristics affected attitudes towards the sites. Non-utilitarian aspects of product use are only addressed by the perceived aesthetic quality of the Web sites. In this sense, the model is very specific, because it demonstrates the effect of perceived aesthetics on emotional responses and not the effect of the perceptions of non-utilitarian attributes in general. Comparing this model to the UX model, environmental stimuli roughly correspond to the level of perceived product character in UX. Aesthetics in general may be considered as a subcategory of hedonic attributes (or an antecedent of perceived hedonic quality), while usability corresponds to pragmatic attributes. Attitudes are consequences of the interaction. In this regard, Porat and Tractinsky's model can be viewed as a specific case of a general interaction-experience model for a specific application domain (i.e., Internet shopping sites), that has a merit of being more detailed with regards to aesthetic determinants and affective reactions.

2.6.3. Components of user-experience model

In a series of experiments, Thüring and Mahlke (2007) selectively manipulated system properties (usability and aesthetics) of interactive products (layouts of simulated mobile-telephone interfaces and portable audio players) and measured changes in emotional, behavioural and evaluative responses. They summarised their findings in a model of components of user-experience (CUE model) that aims to integrate the most important aspects of human-computer interaction. The model regards interaction experience as a compound of three basic elements: (1) the perception of instrumental qualities, (2) the perception of non-instrumental qualities and (3) a particular user's emotional responses to a system (see Figure 2.9).

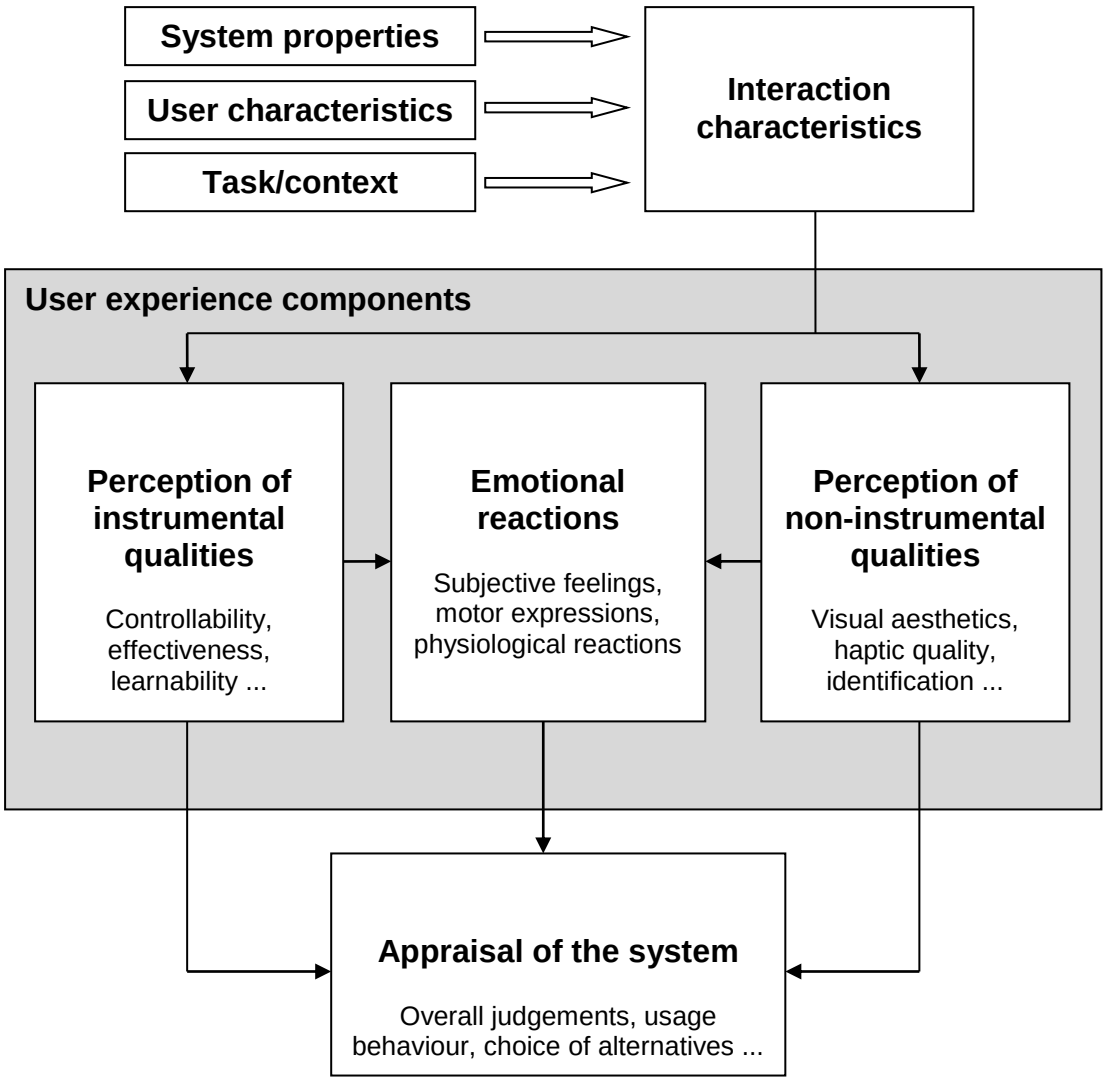


Figure 2.9. The components of user experience model (Thüring & Mahlke, 2007).

Instrumental qualities in the CUE model concern the usefulness and ease of use of a particular system and correspond to pragmatic attributes in the UX model. Non-instrumental qualities in the CUE model concern the 'look and feel' of a particular system (visual aesthetics or haptic quality) and correspond to hedonic attributes in the UX model. This category incorporates aspects that seem to be important to users but are not connected to their performance with a system.

The CUE model claims that the perception of instrumental and non-instrumental system qualities exhibit emotional responses that accompany the interaction process. Emotions are characterised as “episodes of subjective feelings accompanied by specific physiological reactions and expressive behaviour” (Thüring & Mahlke, 2007, p. 263). Mahlke and Minge (2008) emphasize the importance of measuring emotions using multiple components, that is, the notion that emotions are composed of subjective feelings, motor expressions, physiological activations, cognitive appraisals and behavioural tendencies. Mahlke and Minge (2008) used of a wide range of methods for measuring emotions in the context of interaction-experience.

In the CUE model, the characteristics of a particular human-system interaction are composed of three groups of variables: system properties, user-characteristics and task/context characteristics. System properties include, for example, functionality and interface design. This category corresponds to the artefact component in the PAT framework and product features in the user-experience model. Through the system-properties component, the CUE model aims to connect designable product characteristics to the perception of product characteristics, similar to the UX model.

User-characteristics include the knowledge and skills of users. This category corresponds to the person component of the PAT framework. Several other user characteristics may be considered relevant to interaction experience, which are not included in the CUE model. It is possible that people with certain personality characteristics respond differently to a given interactive product. An example to this is *need for cognition*, defined as “an individual’s tendency to engage in and enjoy effortful cognitive endeavours” (Amichai-Hamburger et al., 2007, p. 882), which showed to affect people’s Web-site preferences (Kaynar & Amichai-Hamburger,

2008). The variables of *computer playfulness*, *computer self-efficacy* and *computer anxiety*, discussed in correspondence with the TAM model, may be considered as established measures of user-characteristics. The moderating variable *experience* in UTAUT may be considered as a user-characteristic variable, along with *gender* and *age*, which are not explicitly present in the conceptual representation of the CUE model.

Task/context characteristics in part correspond to the task component of the PAT framework. However, the authors do not specify what type of variables does the task/context component include. From the previously discussed models and literature, the constructs of *voluntariness of use* (TAM) and *mode of use* (UX model) can be considered as plausible determinants of the task/context component.

The CUE model has the merit of addressing person, artefact and task factors as determinants of interaction characteristics in its conceptual framework, which determine the quality of a user's experience. However, the three studies that formed the basis of the model did not include manipulation or measurement of person or task/artefact characteristics. Interaction experience is divided into three components, two of which correspond to hedonic qualities and pragmatic qualities in the UX model. Non-instrumental qualities contain visual aesthetics, similarly to the environmental-psychology model, although it is measured using a different approach (Thüring & Mahlke, 2007). Non-instrumental qualities also include identification and haptic quality, indicating that there are other categories of instrumental qualities that were not addressed directly in the experiments conducted by the authors. In Figure 2.9, three dots in the perception of instrumental and non-instrumental qualities boxes indicate that other categories and measures of these components may be included.

The CUE model aims to be a general model of interaction experience and it incorporates all principal components identified in the PAT framework.

2.6.4. Model of user-interface quality assessment

The model of user-interface quality assessment (Hartmann et al., 2008) aims to describe users' decision-making processes in the assessment of interactive interfaces. The model is based on Adaptive Decision Making theory (Payne et al., 1993). It posits that users' decision-making and quality judgements are adaptive and they are dependent on task- and context characteristics, as well as on the users' previous experiences.

Hartmann and colleagues (2008) used two different versions (menu-based and metaphor-based) of two Web sites (Sutcliffe & De Angeli, 2005; De Angeli et al., 2006) with exactly the same information content and assessed participants' judgements regarding usability, aesthetics (classical and expressive), information and service quality, engagement and overall preference. Although the expressively aesthetic metaphor-based versions of the sites received lower usability ratings, they were consistently preferred by the participants over the classically aesthetic menu-based versions, due to increased engagement and pleasure experienced during interaction. Usability and expressive aesthetics were the most important predictors of users' overall preferences. Participants' preferences for design style changed according to the context of use (serious or leisure) and the target populations (children or mature). For serious use and mature, more experienced target populations, the menu-based design was preferred. The metaphor-based versions were deemed more appropriate by the participants for children and for leisure purposes. The results demonstrated that users made a trade-off between usability

and aesthetics, and they adapted their decision-making processes according to task/context characteristics. However, context of use and person characteristics were not manipulated or measures in the studies; instead, the participants reported their preference judgements according to hypothetical use-scenarios. Another finding of the experiments was that although information content was identical for both versions of the Web sites, the menu-based versions were rated more favourably for information and service quality, probably because the menu-based interfaces facilitated access to content.

In a study that applied three departmental Web sites of a university, Hartmann et al. (2007) found that the importance of design aesthetics and content quality in overall preference was subject to scenario of use: participants reported the preference of a Web site with higher quality of content for PhD studentship (serious use), and a site with higher perceived aesthetics for a summer internship (less serious use). Additionally, the ranking of the three sites based on perceived aesthetics differed significantly between technical and design students, which demonstrates differences in aesthetics preferences due to differences in users' background.

Another experiment addressed the effects of customisation and aesthetics on overall preference judgements, using two aesthetically different versions of mobile telephone interfaces (Hartmann et al., 2008). Both aesthetically different interfaces provided access to the same online news content from a pre-defined list of online magazines in the generic condition. In the content-fit condition, participants were asked to select a set of magazines they wished to access from a list (customisation). The customised versions were perceived as more usable than the generic versions, and also received higher look-and-feel ratings. All participants preferred the customised

interface over the generic one and the majority of participants preferred the more aesthetically pleasing interface version. However, when the participants were asked to make a trade-off choice between customisation and aesthetics, a 50:50 split appeared between people who would sacrifice customisation for aesthetics and those who would sacrifice aesthetics for customisation. Those who preferred customisation over aesthetics judged the effect of customisation more positively. Those who preferred aesthetics over customisation rated the less aesthetically pleasing interface significantly worse than the others. The results showed that the relative importance of customisation and aesthetics in overall preference ratings was affected by individual preferences.

The findings of the experiments are summarised in a model of user-interface quality assessment (see Figure 2.10). The model proposes three stages in users' judgement following the sequence of experience from initial encounter with the artefact to actual interaction. On initial encounter, users assess the application according to their goals and the task domain. Next, decision-making criteria are selected according to their goals and tasks. For example, if a particular user accesses a news site to search for a particular piece of information, he/she might prefer classical aesthetics with clear navigation structure and high level of usability over an expressive interface design. The experience of use will be judged against these selected criteria. If a criterion is selected, it will influence the overall evaluation of the artefact. Decision-making for user-interface quality assessment is an iterative process: users constantly modify their opinion as their experience progresses. Although the causal sequence between the assessment of the relevant criteria and the overall judgement is not specified in the model, the authors suggest that overall preference is a consequence of users' experience in light of the selected criteria. An

alternative causal sequence would be that users justify their overall preferences by selecting and evaluating various criteria.

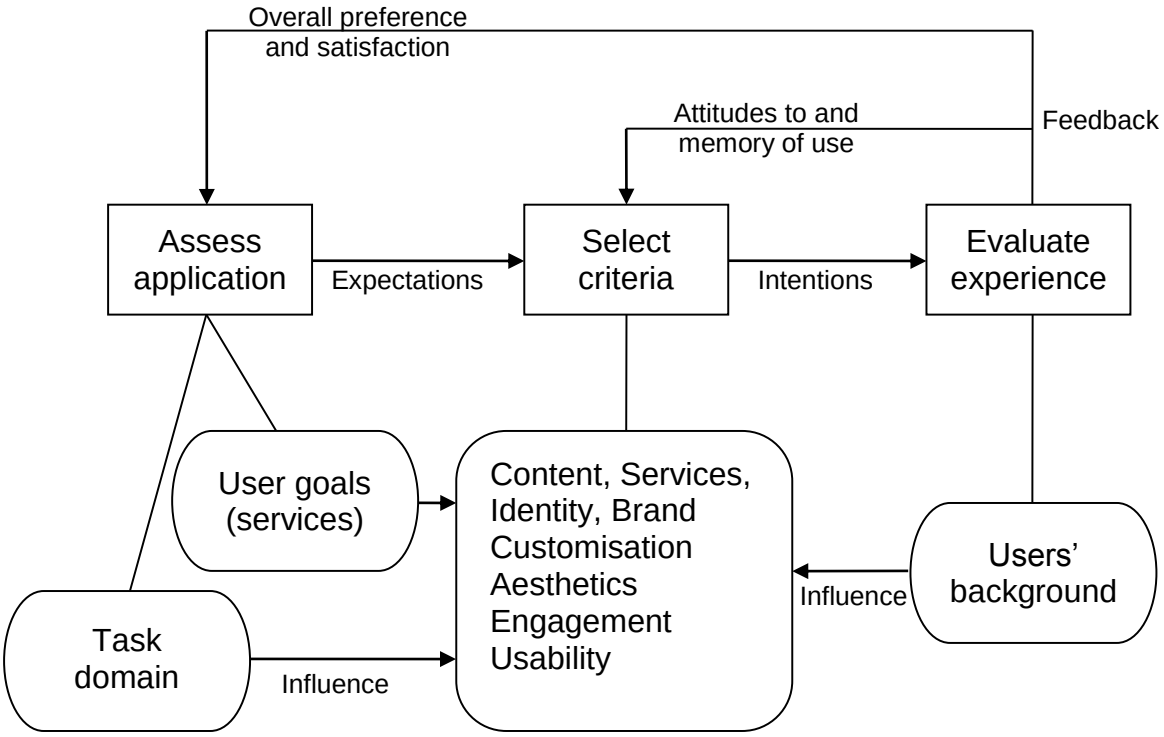


Figure 2.10. The model of users' decision-making process for user-interface quality assessment (Hartmann et al., 2008).

The model of user-interface quality assessment places emphasis on the contextual and task-related factors of interaction and aims to account for changes in assessments of an artefact in distinct situations. Users' background may be considered as a representation of person characteristics in the model, although it was not directly related to the participants' interaction experience in the studies (i.e., participants merely recommended the use of particular design versions to specific target populations). The momentary, subjective, use-centred and evaluative nature of experiences is made explicit. Perceived product characteristics, such as *usability* and *aesthetics*, are not separated into instrumental/non-instrumental or

pragmatic/hedonic categories as in the UX and CUE models; instead, they are presented in one category with product characteristics (e.g., *content* and *customisation*) and *engagement*. The importance of the variables which serve as evaluation criteria in the formulation of experience are subject to trade-offs (usability versus aesthetics and aesthetics versus customisation). The model describes the quality assessment of artefacts as a dynamic, highly context-dependent and iterative process.

2.7. Discussion of models of interaction experience

Fourteen publications were identified in relation to the models of interaction experience presented in this chapter, containing 21 studies. In order to assess the comprehensiveness of the research work using the models in terms of the PAT framework, each study was reviewed to identify person, artefact and task/context variables. Additionally, measures of experience, interaction outcomes and objective measures were collected from the studies. The collected measures were assessed in terms of time of measurement: applied before interaction, during interaction or after interaction. The artefacts used in the studies and the tasks undertaken by the participants were also listed to assess the areas of application and range of tasks the models were applied to. To provide a transparent overview of the studies and for brevity, only a simplified review is presented in Table 2.1. The complete summary of the studies containing the names and dimensions of variables is presented in Appendix 2.1.

Table 2.1. A summary of variables in studies related to the interaction-experience models.

Source	Person	Artefact	Task/context	Experience	Objective
User-experience model					
Hassenzahl et al. (2002)		✓	✓		
Hassenzahl (2004)[a]		✓			
Hassenzahl (2004)[b]		✓		✓	
Hassenzahl and Ullrich (2007)			✓	✓	✓
Hassenzahl (2008)				✓	
Hassenzahl et al. (2010)		✓		✓	
Hassenzahl and Monk (2010)		✓			
Van Schaik et al. (under review)[a]		✓			
Van Schaik et al. (under review)[b]		✓			
Van Schaik et al. (under review)[c]		✓	✓	✓	
Van Schaik and Ling (2008)		✓		✓	✓
Van Schaik and Ling (2009)[a]		✓	✓		
Van Schaik and Ling (2009)[b]		✓	✓	✓	✓
Van Schaik and Ling (2011)		✓	✓	✓	✓
Environmental-psychology model					
Porat & Tractinsky (in press)		✓		✓	
Components of user-experience model					
Thüring and Mahlke (2007)[a]		✓		✓	✓
Thüring and Mahlke (2007)[b]		✓			✓
Thüring and Mahlke (2007)[c]		✓		✓	✓
Model of user-interface quality assessment					
Hartmann et al. (2007)	✓	✓	✓		
Hartmann et al. (2008)[a]	✓	✓	✓		✓
Hartmann et al. (2008)[b]		✓	✓	✓	✓

Note. Bold highlights indicate the experimental manipulation of at least one variable in the category.

[a]Study 1 in the referenced publication. [b]Study 2. [c]Study 3.

2.7.1. User-experience model

The user-experience model divides perceived product characteristics into two distinct categories: pragmatic and hedonic attributes. Usability is addressed in pragmatic attributes in the model. Hedonic attributes emphasise psychological well-being through the fulfilment of be-goals (e.g., stimulation, identification and evocation). The UX model connects the perception of product attributes to product features and links hedonic attributes of interactive products to universal human needs. Positive experience is a result of the perception of hedonic and pragmatic attributes.

Interaction experience in the UX model is characterised as consequences of the interaction: pleasure, satisfaction and the positive appraisal of the product (appeal). However, research using the model typically applied beauty and goodness overall product evaluations as outcome measures of interaction experience. Measures of experience in research included mental effort, spontaneity, perceived challenge, perceived enjoyment, need fulfilment and affect. Mode of use, as a variable of task/context characteristics, was manipulated in five experiments. Artefact characteristics (e.g., complexity and usability) were manipulated in five experiments, and four experiments involved the simultaneous manipulation of mode of use. Research using the model applied a wide variety of artefacts, including various types of information-presenting Web sites (e.g., Wikipedia and university departmental sites), which suggests potential applicability of variables and concepts in the model for the purpose of the current thesis. Objective measures of task performance were also applied in the related studies. However, person characteristics were not studied in research related to the model.

In summary, the user-experience model received the most empirical support from experimental studies among the models of interaction experience presented in this chapter. Artefact and task/context characteristics were measured and systematically manipulated in research, but person characteristics were not measured. A wide-range of measures of experience were applied in related studies, but measures of interaction outcomes are typically restricted to overall beauty and goodness judgements of interactive artefacts. Finally, although the user-experience model is fundamentally a causative model that aims to account for positive experiences (and judgements of interactive products) based on universal human needs and users' perceptions of artefact qualities, recent research work using the model investigated

the role of general-to-specific inference processes in the formulation of product evaluations (Hassenzahl & Monk, 2010; van Schaik et al., under review).

2.7.2. Environmental-psychology model

Similar to the UX model, the environmental-psychology model incorporates perceptions of instrumental and non-instrumental qualities under the label of environmental stimuli. Usability is directly addressed in the model. Non-instrumental qualities only include perceived aesthetics. Interaction experience is captured in three dimensions of emotional states: pleasure, arousal and control. Attitudes are the consequences of interaction experience. The perceptions of artefact characteristics are described with usability and the two aesthetics dimensions. Person and task/context characteristics are not included in the model. The model is restricted to the application domain of Internet shopping sites, although it seems plausible that it is applicable to Web sites in general. No objective measures and no systematic manipulations of artefact characteristics were involved in the study that formed the basis of the model.

2.7.3. Components of user-experience model

The conceptual framework of the CUE model is the most general of the models of interaction experience discussed in this chapter. It incorporates all factors that are present in the other models discussed here. The CUE model offers a richer account of emotional responses by addressing the measurement of motor expressions and physiological reactions, compared to research related to the UX model and the environmental-psychology model, where only subjective feelings were addressed. Interaction experience is characterised in the CUE model as the interplay between the perception of instrumental qualities, the perception of non-instrumental qualities

and emotional responses. Similar to the environmental-psychology model, the CUE model incorporates aesthetics in the category of non-instrumental qualities. The category of instrumental qualities also incorporates identification, similar to the UX model. The CUE model lists person and task/context factors as determinants of interaction characteristics, although these factors are not elaborated in the model and were not included in measurement in the related studies. However, the artefact-characteristics of usability and aesthetics were systematically manipulated in research. Artefacts used in research were computer simulations of mobile-telephone and portable-audio-player interfaces. Measures of interaction outcomes included overall ratings and preference judgements of artefacts.

2.7.4. Model of user-interface quality assessment

In the model of user-interface quality assessment, pragmatic and hedonic attributes are not separated to different categories as in the UX and CUE models; instead, they are all considered as various decision-making criteria. The salience of these criteria for experience depends on their relevance to the actual goals of users and the characteristics of the task at hand. The quality of the experience is judged against these criteria, and it is conceptualised as *engagement* in measurement. Interaction outcome is conceptualised as overall preference judgements of interactive products, according to different scenarios of use by different target audiences. However, in the study of Sutcliffe and De Angeli (2005) (reported in Hartmann et al., 2008 in relation to the model), target audiences (i.e., user background) and scenario of use were not manipulated, rather they were hypothetically proposed. Therefore, participants of the study indicated overall preferences to hypothetical scenarios, rather than actual situations. Changes in users' goals during interactions are similar to the concept of *mode of use* that was previously introduced in relation to the UX model. Similar to

the research work related to the environmental-psychology model, the perception of aesthetic quality is separated into classical and expressive dimensions. Emotional responses are not addressed in the model, although engagement may be considered as a conative aspect of emotions. The manipulation of artefact characteristics in research related to the model included customisation of a particular interface by the users, and the experimental manipulation of aesthetics and presentation style. An additional variable of task/context characteristics was mobility in Hartmann et al. (2008), where users were asked to complete a reading task using a mobile phone during either standing still or walking. Studies related to the model applied Web sites and mobile telephone interfaces as artefacts.

The model of user-interface quality assessment emphasises the dynamic effects of changing user-goals and task characteristics on interaction experience via the weighting of importance of various perceived product characteristics. While other models of interaction experience presented here follow a specific-to-general causative path in describing the formulation of experiences and perceived product character based on the perceptions of product characteristics, which, in turn, account for overall evaluation or consequences, the model of user-interface quality assessment starts with the users' assessment of a particular artefact according to their goals and tasks, and describes quality judgement as an iterative, highly context-dependent, dynamic process. However, the role of context, conceptualised as mode of use, is tested in relation to the UX model to a greater extent. Furthermore, the CUE model incorporates the variables of all reviewed models in its conceptual framework presented in Figure 2.9. *User-goals* and *task domain* correspond to task/context factors in the CUE model. Most of the various decision-making criteria correspond to variables in different categories of CUE. For example, *content,*

services and *customisation* may be regarded as system properties, *usability* as an instrumental-quality variable, and *aesthetics* as a non-instrumental-quality variable and *engagement* as a conative aspect of emotional responses. Users' background may be considered as a category corresponding to user characteristics. Finally, the outcomes, namely *overall preference* and *satisfaction* correspond to the appraisal of the system in CUE.

2.8. Summary and conclusion

This chapter started with defining interaction experience for the purpose of this thesis. Usability was also defined, because it conceptually overlaps with interaction experience and it is addressed in each interaction-experience model. Models of technology acceptance and adoption were presented to provide a theoretical background in human-computer interaction research, and because these models present variables and empirically supported relationships between variables that are relevant to developing a model of end-users' interaction with news sites. Adaptations of the technology acceptance model to the World Wide Web and to the application area of online news services were also included.

Four models of interaction experience were described and evaluated using the PAT framework to identify aspects of experience that need to be addressed in the formulation of a model of interaction experience with news sites. The evaluation of the models concluded that the user-experience model received the most empirical support and it was applied to the widest range of artefacts among the models. The conceptual framework of the CUE model is broad and general enough to incorporate the aspects of interaction experience addressed by the other models, and interaction characteristics in the model correspond directly to the PAT framework. Interaction

characteristics in the conceptual representation of the UX model (as presented in Figure 2.7) contain only artefact properties under the label of product features. However, research related to the UX model typically addressed task/context characteristics and need fulfilment. The environmental-psychology model contains usability and aesthetics as artefact properties. The model of user-interface quality judgements contains artefact characteristics (e.g., usability, content and aesthetics), task/context characteristics (e.g., scenario of use and mobility) and person characteristics (user background). In general, person characteristics were rarely and only indirectly addressed in research work related to the models.

Interaction-experience components in the CUE model include emotions and perceptions of instrumental and non-instrumental qualities. Pragmatic and hedonic attributes in the UX model correspond to instrumental and non-instrumental qualities. Although emotions are not present in the conceptual framework of the UX model, valence dimensions of affective reactions (positive and negative) were addressed in research using the model. The environmental-psychology model also includes the arousal and dominance dimensions of affective reactions. Research related to the CUE model addressed emotional responses most thoroughly among the models by incorporating the measurement of affective reactions (valence and arousal), emotional responses (cognitive appraisals) and objective measures as correlates of emotions (heart rate, electrodermal activity and electromyographic activity of facial muscles). The model of user-interface quality assessment does not include measures of emotions, and instrumental and non-instrumental attributes are not separated from each other and from artefact characteristics.

Measures of interaction outcomes typically involve goodness and beauty ratings of interactive artefacts in research related to the UX model. The environmental-psychology model includes attitudes toward artefacts as the outcome measure of interaction. Research related to the model of user-interface quality assessment used overall preference as the outcome measure. Finally, research that formed the basis of the CUE model included both overall ratings and preference judgements as measures of interaction outcomes. However, van Schaik and Ling (2011) emphasise the need to also address the acceptance of technologies (e.g., intention to use) in interaction-experience to account for repeat and prolonged use of Web sites (i.e., online loyalty).

Each of the models has a different focus. The environmental-psychology model provides a detailed account of how usability and aesthetics affect the emotional state of users, or at least subjective feelings as a component of emotions. The UX model is more detailed in the description of pragmatic and hedonic qualities, and aims to connect positive experiences to fulfilment of universal human needs to identify the source of experiences. The model of user-interface quality assessment emphasises the role of contextual and task-related factors in the formulation of preference judgements. Finally, the conceptual framework of the CUE model integrates the most aspects of interaction experience present in the other models. Usability is and aesthetics are addressed directly in all four models.

Although the review of models of technology acceptance and interaction experience provided an extensive range of aspects of experience that could be included in the measurement model of interaction experience with news sites, the available research literature applying the models did not include empirical work related to this particular

application area. Therefore, an exploratory study was conducted to identify aspects of interaction experience which are particularly relevant to the application area of online news, in order to guide the development of a measurement model for further studies presented in this thesis. The exploratory study applied protocol analysis on verbal data to derive self-reported aspects of experience of users' everyday interaction with a news Web site. The method selection for the collection of verbal protocols and the exploratory study are presented in the next chapter.

Chapter 3

An exploratory study of users' experiences with a news site[7]

[7] A paper based on this chapter was first submitted to the British HCI (2011) Conference and recently published by the journal *Interacting with Computers* (Aranyi et al., 2012).

Contents

3.1. Overview

In order to complement the literature review, an exploratory study was conducted to gain insight into how users of a particular online news site interact with the site and how they report their experience. A concurrent think-aloud protocol was applied in a laboratory setting to capture participants' verbal comments and news-site use behaviour while interacting with a particular news site, Gazette Live (*http://www.gazettelive.co.uk*). An online questionnaire was used to collect data on how the participants of the think-aloud study and regular users of Gazette Live use the Internet in general and news sites in particular, along with demographic details.

The analysis of users' news-site use and verbalisation of their thoughts when using a particular news site, together with the literature review, provided a basis for the following phases of the research project reported in this thesis. In particular, the results were used to steer the formulation of the construct of interaction experience and the selection of its factors to be included in a measurement model of interaction experience (Pedhazur & Pedhazur-Schmelkin, 1991).

This chapter starts with the description of the think-aloud method and the possible approaches to applying think-aloud protocols. Following the general discussion of think-aloud, the particular method that was used in the exploratory study is presented. The general description of the method and procedure of the study is followed, first, by the analysis of the online questionnaire, then, by the analysis of the think-aloud recordings and, finally, the set of questionnaires completed by the participants after thinking aloud. The findings of the study and its implications for the following research phases are summarised.

3.2. The think-aloud method

3.2.1. A general description of think-aloud protocols

Think-aloud protocols have been used extensively in usability testing and this method is well-established in practice and research within human-computer interaction (e.g., Nielsen, 1993, 2012). A think-aloud protocol can be employed in many ways, leading to various consequences in terms of practical considerations as well as concerns about reliability, validity and the appropriateness of the selected procedure to the given area of interest. In this section, think-aloud is briefly introduced and some practical issues related to the method are discussed in order to demonstrate the rationale of the procedure used in the current study.

In think-aloud sessions, participants are usually asked to carry out tasks while verbalising their thoughts. The verbalisations are recorded and analysed in order to gain information about the participants' thought contents and processes. When the think-aloud method was first introduced, it was used to study human cognition and thought processes (Gilhooly & Green, 1996). In the context of human-computer interaction and usability, this method is generally used in the evaluation of products. The focus is on a particular product rather than on the user of the product. The results of usability tests are used to identify possible usability problems with a product, but they also carry information about test-users' perceptions of the product and important details of user-product interaction.

Thinking-aloud aloud can be considered as an externalization of information passing through working memory (Ericsson & Simon, 1993). In the course of thinking aloud, conscious contents of working memory are externalized in a verbal code so they can be observed and studied. However, there is uncertainty about how informative

verbal protocols collected this way are to shedding light on the full extent of thought processes. The completeness of information collected by think-aloud would be compromised if one assumes that thought processes run in parallel (i.e., the idea of unconscious streams of thought). If there are parallel thought processes in the course of the interaction with an artefact, then even the most detailed verbal reports obtained through think-aloud may only represent a fraction of information. For the purpose of the present research, however, there is no need to further discuss the theories of human thought processes, because it is not expected that the think-aloud method will illuminate every aspect of thought about a given artefact (for a detailed discussion of serial versus parallel thought processes in relation with think-aloud methods, see Gilhooly & Green, 1996).

Here the intention is not to achieve a thorough description of interaction experience with news sites using this particular method. The main assumption guiding this study is that users of an artefact possess valuable information about a given artefact in the form of verbalisable thoughts, and (at least part of) this information can be captured accurately if the verbalisations are recorded in the course of use. Moreover, the focus is on how users report their experience and which elements of the artefact they notice as being important in the interaction process. In this approach, users are regarded as experts in using the artefact. The focus is on the users' interaction experience, and not the artefact at hand or its usability. Every utterance referring to a feature of the artefact can be informative in two ways: it conveys information about the attributes of the artefact and, at the same time, it carries the information that it was selected by the user for some reason. By attracting attention, a particular feature is identified as being part of the experience, at least to some extent, and this is made conscious and explicit. However, the verbalisation of a particular aspect of

interaction does not mean that that aspect is an important component of the experience. For example, a colourful picture in the centre of a Web page may be so conspicuous that it triggers a comment from the majority of participants, without having anything to do with their satisfaction with the site. A negative example may be that something is indeed important for interaction experience, but is so hard or inconvenient to verbalise that participants might not report it. Furthermore, certain features or qualities of features may be of relatively great importance to interaction experience, but so obvious at the same time that participants do not even recognise that they are being influenced by them. It is important to bear in mind that not all impressions evoked by an artefact are conscious, but certain subconscious contents of thought, such as emotional responses or sensory impressions, can become conscious and reported verbally.

Think-aloud protocols involve the recording and analysis of verbalisations produced by participants while a (primary) task is being completed. These protocols are usually applied to gain insight into the way actions and tasks are carried out normally, without the involvement of overt verbalisation, which raises the issue of reactivity: the problem that thinking aloud may in itself distort normal thought processes to some extent. Ericsson and Simon (1993) reviewed a broad range of studies and concluded that in most cases concurrent think-aloud has no significant effect on the quality of performance. There are, however, some instances when the extra workload of verbalising can impair the performance in demanding tasks (e.g., mental arithmetic). Russo et al. (1989) tested the reactivity and veridicality (the extent to which the protocol reflects the underlying process) of concurrent and retrospective think-aloud protocols with four distinct types of task (a verbal anagram task, a numerical gamble task, Raven's progressive matrices as a pictorial task and a mental arithmetic task).

They found a general prolongation of reaction time when thinking-aloud was involved. The accuracy of task completion was compromised in the anagram and arithmetic tasks, left unaffected in the pictorial task and was even enhanced in the gamble task. These findings suggest that, with certain stimuli, the use of verbal protocols may alter the quality of task performance and the sources of reactivity are complex. In retrospective think-aloud, participants carry out tasks silently and comment on their work afterwards on the basis of their own performance. Therefore, the retrospective protocol is free from the problem of reactivity, but the issue of veridicality arises, that is, how accurately the reports reflect the underlying process of task completion. Russo et al. (1989) suggest that retrospective protocols fail to provide a veridical reflection on the process due to forgetting and fabrication.

With regards to the application of think-aloud in human-computer interaction, van den Haak et al. (2004) compared concurrent think-aloud, retrospective think-aloud and constructive interaction protocols in the evaluation of an online library catalogue. During constructive interaction (‘co-discovery learning’), two participants work together on a task and verbalise their thoughts through interacting with each other. Constructive interaction is considered to be a more natural setting to elicit verbal reports from participants. The authors did not find any significant differences between the three methods in terms of the quantity and relevance of usability problems detected, although there were some differences in how the actual usability issues surfaced. They found that the participants in the concurrent think-aloud condition did not perform their tasks slower or less accurately than those in the retrospective condition and the three methods roughly produced the same findings. For practical reasons, the authors made a case for the concurrent think-aloud protocol: it is less time-consuming than the retrospective think-aloud protocol, where

twice the time is required to allow the participants to watch their performance and comment on it, and it requires half as many participants as the constructive interaction protocol.

As demonstrated above, for collecting records of verbal behaviour during a task, the concurrent think-aloud method is but one of (at least) three methods that can be used. (The other two methods are retrospective think-aloud and co-discovery learning.) Concurrent protocols have many advantages though, which can make them theoretically and/or practically more attractive. However, there are also issues regarding the application of the concurrent think-aloud method itself.

According to the theoretical framework proposed by Ericsson and Simon (Ericsson & Simon, 1993), researchers should keep their interactions with participants to a minimum during think-aloud sessions in order to prevent them from interfering with the participants' thought processes. A particular participant's mental processes should not be influenced in any way by the researcher. The only intervention the researcher is allowed is the phrase 'keep talking', in order to remind the participant of the task in case he/she falls silent for a given period of time (typically after 15-20 seconds). Participants are usually told in the instructions to talk as if they were alone in the room. In addition, during a session the researcher remains out of the participant's line of sight, excluding him/herself from the interaction of the participant and the artefact. The goal of the setting is to elicit undirected, undisturbed and constant verbal reports from the participant. If these rules are strictly applied, only spontaneous, uninfluenced verbalisations count as data. Verbalisations that are influenced by interventions from the researcher are results of an interaction between the participant and the researcher. In such cases, the researcher may introduce new

contents to the participant's working memory; therefore, such verbalisations should not be treated as reliable and should be discarded as data. Adhering to this protocol, the researcher does not need to worry about influencing the verbal utterances of participants. The participants' verbal utterances can be treated as representing their mental processes; thus validity is preserved. The strict application of the protocol has some drawbacks though. Because the participants are basically 'left alone' after the instructions, as there is no interaction other than the prompt to keep talking allowed between the researcher and the participant until the end of the session, researchers may fear losing valuable data if participants are not probed for more information when they arrive at a point of particular interest. If a participant is stuck in a loop or violates the instructions, the researcher has no means to intervene without violating the rules of the protocol.

Boren and Ramey (2000) reviewed how the actual practice of applying think-aloud protocols in usability research diverges from Ericsson and Simon's prescriptions for using the method, which is the most often cited theoretical background for usability studies applying think-aloud protocols, and proposed an alternative method. Boren and Ramey point out that practitioners do not apply think-aloud protocols uniformly, making comparisons between results from different studies difficult. The incongruity in the application of the method is mainly attributed to the difficulties of applying Ericsson and Simon's cognition-focused framework to the domains of usability and human-computer interaction, which are more task- and system-oriented, and to the technological contingencies of usability sessions (e.g. system crashes and bugs) requiring intervention, thus interrupting the participants' monologues. Boren and Ramey criticise Ericsson and Simon's protocol for creating a rather unnatural setting in an attempt to preserve validity by minimising person-to-person interactions. In

their speech-communication approach to conducting think-aloud protocols, a particular participant's verbalisations are not monologues, rather parts of a dialogue with the researcher, although the participant talks most of the time.

The speech-communication protocol entitles the researcher to take part in this rather asymmetrical conversation, where he/she does not try to 'disappear' from the participant, but minimises his/her contribution to the content and restricts utterances to feedback and reminders. This way, the problem of unnatural interaction between the researcher and the participant during the think-aloud session can be reduced. Participants anticipate an active and responsive listener while speaking, and their speech is acknowledged by the researcher using various communicative tokens. Participants do not feel like 'talking to the wall', because they are constantly reassured about the interest of the listener. The researcher is allowed to make acknowledgements as feedback, using various acknowledgement tokens, such as 'okay', 'yes', or conversational grunts ('mm-hmm'). These tokens may also be used as reminders when a participant falls silent for more than 15-20 seconds. Furthermore, the researcher is allowed to repeat a single word with appropriate inclination to trigger clarification. According to Ericsson and Simon's framework, such interventions break the natural flow of thought of participants, resulting in less reliable data. However, when the focus is on an artefact (e.g., the usability of a software application) rather than on the participants' mental processes, such interventions can be considered reasonable and beneficial in acquiring more relevant data on the subject of interest.

As Krahmer and Ummelen (2004) point out, there are three possibilities for researchers intending to use a think-aloud protocol. First, they can adhere strictly to

Ericsson and Simon's prescriptions for using the think-aloud method, because it can yield higher validity and reliability, but with the possibility that it may lead to fewer and less informative data. Second, a modification of Ericsson and Simon's protocol can be applied with several unspecified deviations. This 'simplified think-aloud', however, can lead to various problems of validity and reliability. Third, researchers can use the speech-communication protocol, which is expected to yield more relevant data by allowing proactive interventions to elicit more information at the expense of some possible concerns regarding validity. In their study, Krahmer and Ummelen (2004) compared Ericsson and Simon's strict protocol with the speech-communication protocol proposed by Boren and Ramey (2000) by conducting usability testing on an artistic Web site with highly non-conventional navigation and presentational style. No significant differences were found between the two methods in the numbers of usability problems detected, nor in the amount of verbal utterances of participants. Although Boren and Ramey's protocol did not show any advantages with regards to the detection of usability problems, participants under this condition were less likely to get lost in the Web site and were engaged in a more natural interaction with the researcher. These findings imply that when deciding between the two think-aloud approaches, there is a trade-off to be considered. On the one hand, Ericsson and Simon's protocol offers higher validity, sacrificing the extent to which the setting appears natural. On the other hand, Boren and Ramey's protocol offers a more natural interaction between a particular participant and the researcher, raising the important issue of the participant's behaviour being influenced by the researcher's interventions, thus compromising the validity of verbal reports. Either way, the outcome is expected to be approximately the same.

3.2.2. The think-aloud protocol of the exploratory study

Given the primarily exploratory nature of this study and the aim of capturing everyday users' spontaneous reports of experience (and not accounts regarding the usability of a particular Web site), it was concluded that the least amount of researcher-participation is favoured, even if the consequence is that potentially interesting details are left unexplored because of the lack of prompting or directing the participant towards areas of interest. However, as mentioned among the disadvantages of Ericsson and Simon's strict think-aloud protocol, the danger of losing valuable data had to be accounted for; otherwise, the value of the study could be seriously reduced. Instead of prompting participants while thinking aloud to encourage them to give rich and detailed verbal reports of their experience, a video demonstration was presented between the verbal think-aloud instructions and the practice task preceding data collection. This two-minute-long video demonstration featured a person who demonstrated how to think-aloud while completing a task. The demonstrator's task was completely unrelated to the one the participants had to perform: he divided an apple into four segments using a pocket knife. During task completion, he gave a rich and detailed description of what he was doing and why, along with the perceivable details of the objects he was using. It was assumed that, without a demonstration, the participants would fail to express non-instrumental elements of their experience; in other words, those elements that are not directly associated with the tasks they are completing and the goals they are pursuing (e.g., the look-and-feel of the site and aesthetic impressions). The abundance of such utterances in the video demonstration was expected to prime participants to incorporate these elements in their verbal reports. The video demonstration served a

further purpose: by demonstrating how to think-aloud, it aimed to make the instructions clear for the participants.

Although it is technologically possible to conduct think-aloud protocols in remote settings (Bastien, 2010), a laboratory setting was chosen for the current study. Remote think-aloud has the advantage that participants are not required to come to the laboratory (they may be at home for instance), but remote settings also require teleconferencing and application-sharing tools in order to record verbal and screen-capture data and maintain communication. There is also an increased chance of technical problems to occur during the remote think-aloud sessions. Furthermore, remote think-aloud does not allow for controlling several aspects of the interaction. For example, the Internet connection speed and the characteristics of computers (e.g., screen size and Web browser) may vary from participant to participant, resulting in different experience characteristics for each interaction observed.

The following think-aloud instructions were devised and read out loud by the researcher to the participants:

“In this experiment, we are interested in what you think while using a Web site. I am going to ask you to THINK ALOUD as you explore the site. What I mean by ‘think aloud’ is that I want you to tell me EVERYTHING you are thinking from the time you first see the Web site until you are finished exploring the site or I tell you to stop. I would like you to talk aloud CONSTANTLY during your use of the site. I don’t want you to plan out what you say or try to explain to me what you are saying. Just act as if you are alone in the room speaking to yourself. It is important that you keep talking. If you are silent for some time, I will ask you to

talk. Please try to speak as clearly as possible, as I will be recording you as you speak. Do you understand what I want you to do?"

Note that these instructions made a distinction between think-aloud and explanation. According to Ericsson and Simon's protocol, participants were instructed to act as if they were alone in the room, discouraging the interaction between the participant and the researcher. Participants were told that they would be reminded to talk if they fell silent.

After reading the instructions, participants were given the opportunity to ask questions about how to think aloud. The researcher summarised and clarified the main points of the instructions if this was necessary. The video demonstration was shown on the computer screen[8] and was followed by a short practice, in which the participant used an unrelated Web site (*http://www.nhs.uk*) to ensure he/she understood the instructions. The practice exercise was included in accordance with Ericsson and Simons' suggestion that participants should be given opportunity to practice thinking aloud before the test session. Finally, the following text was read out loud by the experimenter:

"We shall start now. Please use the site as you normally would while you think aloud. Please do not leave the Gazette site. You are free to leave the home page, but please don't navigate away from Gazette. If you navigate away, I'll ask you to go back. Please carry on working until you are told to stop."

[8] The demonstration video had been recorded in the same interview room in which the sessions took place.

After the final instructions, the screen-capture software was started and the Gazette Live site was loaded in a new browser window by the researcher, who then sat back, out of the participant's line of sight. The participant then started to use the site while thinking aloud.

The think-aloud protocol used in the study was piloted with three undergraduate university students who had no experience with this method before. The pilot participants reported that they understood the instructions, which was supported by the researcher's observations of the participants' practice sessions. The participants reported think-aloud to be rather 'strange' at first, but they got used to doing it after a short time. These comments underline the importance of including a practice task before the main task. The demonstration video was reported to be useful by the participants in clarifying the instructions. Furthermore, the video demonstration noticeably helped 'loosening the mood' of participants, who felt a bit tense at first because of the formal experimental setting and environment.

3.3. Method

3.3.1. Design

In order to record participants' use of online news sites and their verbal comments regarding their experience, a concurrent think-aloud study was carried out in a laboratory setting at Teesside University. Participants were recruited among frequent users of a local online news medium (Gazette Live) and they were asked to interact freely with the site under think-aloud instructions. Another group of participants was included in the study, composed of people who were regular readers of online news sites, but did not use Gazette Live. The participants' verbal behaviour was recorded, along with their online activity using screen-capture software. Following the main

think-aloud task, the participants completed a set of interaction-experience questionnaires.

3.3.2. Participants

Participants were recruited through advertisements in local newspapers (Evening Gazette, Herald and Post) and on the Gazette Live news site to reach audiences who used the site regularly. These advertisements invited people to fill out an online questionnaire to collect data, including demographic details, Internet use behaviour and news-site use behaviour. At the end of the questionnaire, people who were interested in participating in a practical session at the university were asked to provide their e-mail addresses. Those who did were invited to take part in the think-aloud study and were offered a monetary compensation for their time and effort. Another group of participants, who did not use Gazette Live but were regular readers of online news sites, were recruited among Teesside University staff and postgraduate students through the university's intranet. This group filled out the same online questionnaire. All participants had to be over 18 years and had to be fluent in English to be eligible for the study. Seventy-seven gave a valid response to the online questionnaire (43 male, 34 female; mean age = 37.49 years, *SD* = 13.95; 62 regular users of Gazette Live and 15 non-Gazette users). Twenty-five participants attended the think-aloud sessions (10 Gazette Live readers and 15 regular users of other online news sites; 13 male, 12 female; mean age = 32.00 years, *SD* = 10.84). The think-aloud participants received £8 compensation for their time and effort.

3.3.3. Materials and equipment

The items of the online questionnaire are presented in Appendix 3.1. The think-aloud sessions were run individually in an interview room at the university, using the same

computer (OS: MS XP Professional; Web browser: MS Internet Explorer 7.0; Monitor: 19'' TFT, 1280x1024 resolution, 60Hz refresh rate; screen-capture software: Camtasia Studio 3.0). Participants used the National Health Service (NHS) Web site (*http://www.nhs.uk*) for practicing thinking aloud and the Gazette Live news site (*http://gazettelive.co.uk*) was used for the main think-aloud task. A 10-item perceived aesthetics questionnaire (Lavie & Tractinsky, 2004) was used to examine the participants' perceptions of Web site aesthetics on two dimensions: classical and expressive aesthetics. A 9-item intensity of flow questionnaire (Davis & Wiedenbeck, 2001) was used to measure the extent to which participants felt engaged in the interaction with the site. A 3-item questionnaire was adopted from Sun and Zhang (2008) to measure perceived enjoyment as an affect dimension of the participants' experience. Perceived enjoyment was defined in this context as the extent to which the activity of exploring the Gazette Live site was perceived to be enjoyable in its own right, apart from any performance consequences that may be anticipated. A 10-item version of the AttrakDiff2 questionnaire (Hassenzahl & Monk, 2010) was used to measure the participants' perceptions on the site's pragmatic and hedonic quality (four items each), along with two overall evaluative judgement items regarding goodness and beauty. The items of the interaction-experience questionnaires (with the participant information sheet and consent form used in the study) are presented in Appendix 3.2.

3.3.4. Procedure

The think-aloud sessions were run individually in a laboratory at Teesside University. First, each participant read the information sheet and gave their written consent to participating. Following this, the think-aloud instructions were read out loud by the researcher. After the participant confirmed his/her understanding of the instructions,

the two-minute long demonstration video was presented on the computer screen. After the demonstration video, participants were asked to practise thinking aloud using the NHS Web site. Following the practice, the researcher loaded the Gazette Live site in a new browser window and started the screen capture and audio recording, handed the mouse over to a participant and sat back, out of the participant's line of sight. The participant started to browse the site freely, without any set tasks while thinking aloud. The researcher and the participant did not interact during the recording session other than the short, non-intrusive 'keep talking' reminders uttered by the researcher when necessary. After 10 minutes, the researcher thanked the participant and stopped the recording. Following the think-aloud task, participants were asked to complete a series of questionnaires measuring various aspects of interaction experience. Finally, participants received their compensation and were thanked for participation. The test session took approximately 30 minutes per participant.

3.4. Analysis of the online questionnaire

3.4.1. Demographics

Demographics data were collected with the online questionnaire to provide a general description of the sample. A summary is presented in Table 3.1.

The majority of the respondents (*n* = 62) were regular users of Gazette. The 15 non-Gazette users who participated in the think-aloud sessions were also asked to fill out the online questionnaire. Their results are presented separately. Two participants were excluded, because their age did not reach the minimum 18 years. Altogether, 77 gave a valid response to the online questionnaire (43 male, 34 female). Their average age was 37.49 years (*SD* = 13.95). Non-Gazette readers had a lower

average age (*M* = 30.13 years, *SD* = 8.90), perhaps because most of the non-Gazette readers were graduate students who were recruited at the university. The age-range of participants was rather wide: the youngest respondent was 18 and the oldest was 69 years old. Over half of the sample held a degree or higher degree. However, of the non-Gazette readers, 60% held a higher degree. Only 3% of the overall sample had no qualification. Regarding occupational status, the majority of respondents were employed (60%) or students (14%) and 8% of the respondents were retired. A summary of the demographic data of think-aloud participants is presented separately in Table 3.2.

Table 3.1. Demographic statistics of the overall sample.

Age	Mean	Minimum	Maximum	*SD*	*N*
Gazette reader	39.27	18	69	14.41	62
non-Gazette reader	30.13	22	51	8.90	15
Overall	37.49	18	69	13.95	77

Gender	Male	Female
Gazette reader	56.5% (35)[a]	43.5% (27)
non-Gazette reader	53.3% (8)	46.7% (7)
Overall	55.8% (43)	44.2% (34)

Highest Level of education	Gazette	Non-Gazette	Overall
O level/GCSE or equivalent	12.9% (8)		10.4% (8)
NVQ or equivalent	11.3% (7)		9.1% (7)
A level or equivalent	16.1% (10)	6.7% (1)	14.3% (11)
TROCN/Btec Dip	4.8% (3)		3.9% (3)
Degree	25.8% (16)	26.7% (4)	26.0% (20)
Higher degree	21.0% (13)	60% (9)	28.6% (22)
No qualification	3.2% (2)		2.6% (2)
Other	4.8% (3)	6.7% (1)	5.2% (4)

Occupation	Gazette	Non-Gazette	Overall
Retired	9.7% (6)		7.8% (6)
Employed	59.7% (37)	60.0% (9)	59.7% (46)
Self-employed	1.6% (1)		1.3% (1)
Student	9.7% (6)	33.0% (5)	14.3% (11)
Full-time homemaker	3.2% (2)		2.6% (2)
Unemployed/looking for work	4.8% (3)	6.7% (1)	5.2% (4)
Unfit to work	8.1% (5)		6.5% (5)
No response/rather not say	1.6% (1)		1.3% (1)
Other	1.6% (1)		1.3% (1)

[a] *n*

Note that all non-Gazette reader participants of the overall sample participated in the think-aloud study, therefore the values under the 'non-Gazette' headings in Table 3.2 are the same as in Table 3.1. Twenty-five people participated in the think-aloud study (13 male and 12 female) with an average age of 32 years (*SD* = 10.84). The majority of think-aloud participants held a degree or a higher degree (24% and 60%, respectively) and most were employed or students (60% and 24%, respectively).

Table 3.2. Demographic statistics of think-aloud participants.

Age	Mean	Minimum	Maximum	*SD*	*N*
Gazette reader	34.80	24	60	13.26	10
non-Gazette reader	30.13	22	51	8.90	15
Overall	32.00	22	60	10.84	25

Gender	Male	Female
Gazette reader	50.0% (5)[a]	50.0% (5)
non-Gazette reader	53.3% (8)	46.7% (7)
Overall	52.0% (13)	48.0% (12)

Highest Level of education	Gazette	Non-Gazette	Overall
O level/GCSE or equivalent	10.0% (1)		4.0% (1)
NVQ or equivalent			0.0% (0)
A level or equivalent	10.0% (1)	6.7% (1)	8.0% (2)
TROCN/Btec Dip			0.0% (0)
Degree	20.0% (2)	26.7% (4)	24.0% (6)
Higher degree	60.0% (6)	60.0% (9)	60.0% (15)
No qualification			0.0% (0)
Other		6.7% (1)	4.0% (1)

Occupation	Gazette	Non-Gazette	Overall
Retired			0.0% (0)
Employed	60.0% (6)	60.0% (9)	60.0% (15)
Self-employed			0.0% (0)
Student	10.0% (1)	33.0% (5)	24.0% (6)
Full-time homemaker			0.0% (0)
Unemployed/looking for work	10.0% (1)	6.7% (1)	8.0% (2)
Unfit to work			0.0% (0)
No response/rather not say	10.0% (1)		4.0% (1)
Other	10.0% (1)		4.0% (1)

[a] *n*

3.4.2. Internet-use behaviour

Participants' Internet-use behaviour was assessed using the online questionnaire in order to gain insight to their Internet use, knowledge and habits. A summary is presented in Table 3.3.

Table 3.3. Descriptive statistics of Internet use behaviour.

Internet use (years)	"How long have you been using the Internet?"				
	Mean	Minimum	Maximum	*SD*	*N*
Gazette reader	10.81	2	20	3.69	62
non-Gazette reader	10.33	6	14	1.92	15
Overall	10.71	2	20	3.41	77

Internet connection at home	"Do you have Internet connection at home?" (Yes)	
Gazette reader	Non-Gazette reader	Overall
96.8% (60)[a]	86.7% (13)	94.8% (73)

Access to Internet	"Where do you access the Internet? (Select all that apply.)"		
	Gazette	Non-Gazette	Overall
Home	96.8% (60)	86.7% (13)	94.8% (73)
Work	53.2% (33)	93.3% (14)	61.0% (47)
Library	4.8% (3)	66.7% (10)	16.9% (13)
Internet café	3.2% (2)	20.0% (3)	6.5% (5)
Other	14.5% (9)	20.0% (3)	15.6% (12)

Devices used	"What kind of devices do you use to access the Internet? (Select all that apply.)"		
	Gazette	Non-Gazette	Overall
Desktop PC	90.3% (56)	100% (15)	92.2% (71)
Laptop	62.9% (39)	93.3% (14)	68.8% (53)
PDA	1.6% (1)		1.3% (1)
Mobile Phone	43.5% (27)	60% (9)	46.8% (36)
Other	3.2% (2)		2.6% (2)

Internet use frequency	"How often do you log on to the Internet?"		
	Gazette	Non-Gazette	Overall
2-3 times a week	1.6% (1)		1.3% (1)
4-6 times a week	6.5% (4)		5.2% (4)
Daily	22.6% (14)	13.3% (2)	20.8% (16)
2-3 times a day	17.7% (11)	6.7% (1)	15.6% (12)
More than 3 times a day	51.6% (32)	80.0% (12)	57.1% (44)

Internet use duration	"Once on line, how much time do you spend on the Internet on average?"		
	Gazette	Non-Gazette	Overall
6-10 minutes	3.2% (2)	13.3% (2)	5.2% (4)
About 15 minutes	14.5% (9)	6.7% (1)	13.0% (10)
About 30 minutes	12.9% (8)	20.0% (3)	14.3% (11)
About 45 minutes	4.8% (3)	6.7% (1)	5.2% (4)
About 1 hour	27.4% (17)	6.7% (1)	23.4% (18)
Several hours	37.1% (23)	46.7% (7)	39.0% (30)

Online application use	"What do you use the Internet for? (Select all that apply.)"		
	Gazette	Non-Gazette	Overall
E-mail	93.5% (58)	100% (15)	94.8% (73)
Reading news	96.8% (60)	100% (15)	97.4% (75)
Buying products or services	88.7% (55)	93.9% (14)	89.6% (69)
Social networking	69.4% (43)	86.7% (13)	72.7% (56)
Visiting chat rooms, forums	19.4% (12)	26.7% (4)	20.8% (16)
Managing bank accounts	67.7% (42)	60% (9)	66.2% (51)
Searching for work-related or study-related information	61.3% (38)	100% (15)	68.8% (53)
Education/training	41.9% (26)	60% (9)	45.5% (35)
Surfing	74.2% (46)	60% (9)	71.4% (55)
Downloading or file sharing	27.4% (17)	46.7% (7)	31.2% (24)
Working	43.5% (27)	86.7% (13)	51.9% (40)

[a] *n*

On average, participants had been using the Internet for 10.71 years (*SD* = 3.41 years, range: 2-20 years). Most participants (95%) had access to the Internet in their homes. The majority of participants accessed the Internet at home (95%) and at work (61%). Most participants used desktop (95%) and laptop (69%) computers to access the Internet. Mobile telephone access was also prevalent (47%). Seventy-two participants (94%) logged on to the Internet at least once a day or more. For the assessment of how often people access the Internet, a *weekly* response category was also available in the online questionnaire, but no participant indicated less than 2-3 times access per week. Almost two thirds of the participants (62%) reported to use the Internet for one hour or more per day. For the assessment of Internet use duration, a *1-5 minutes* response category was also available, but no participant indicated less than 6-10 minutes spent on the Internet daily.

The most-frequent online activity was reading news (97%), followed by e-mail (95%) and buying products or services (90%). For reading news, however, 100% was expected, because the recruitment addressed regular news-site readers. The other online activities listed in the questionnaire in the order of their salience were social networking (e.g., using Facebook and Youtube), browsing the Web ('surfing'), searching for work-related or study-related information, managing bank accounts, working, education/training, downloading or file sharing and visiting chat rooms, forums. Seventy participants (91%) were engaged in five or more of the eleven online activities listed in the questionnaire. These results suggest that the participants were experienced Internet users and being connected to the Web was a part of their everyday lives.

In the online questionnaire, participants were asked to indicate their five most-visited Web sites. The aim of this was to see how salient news site use was and what the ranking of news sites was among most-frequently visited Web sites.

Sixty-two out of 77 (81%) reported at least one news site in the top-five most visited Web sites, and 26 (34%) reported a news site as the most-visited. On average, participants reported using 1.57 news sites (*SD* = 1.21, range: 0-4) in their top five. However, 49% of the participants reported using two or more news sites in their top five.

Among Gazette users (*n* = 62), no participant reported Gazette as the most-frequently visited Web site. However, 16% (*n* = 10) of Gazette users reported BBC and 19% (*n* = 12) reported some other news site as the most-visited Web site. Seven Gazette users (11%) reported Gazette as their second-most visited Web site, 7% (*n* = 4) reported it as third, 16% (*n* = 10) as fourth and 7% (*n* = 4) as fifth. Overall, 40% (*n* = 25) of Gazette users reported Gazette among their five most-visited Web sites. This was second to the BBC site, which was reported by 42% (*n* = 26) of Gazette users. Forty-nine Gazette users (79%) reported a news site other than Gazette in their top five.

In summary, the majority of participants reported news sites among their most-visited Web sites. Half of the participants reported two or more news sites in their top five. One third of participants reported a news site as their most frequently visited Web site. However, these numbers may not represent the popularity of news-site use among other Web sites in the general population, because participants had been recruited among people who were regular users of Gazette, a local news site, and who were regular users of online news sites. Furthermore, no Gazette user indicated

Gazette as most-frequently visited and less than half reported Gazette among their five most-visited Web sites.

3.4.3. News-site use behaviour

The news-site use behaviour of participants was also assessed in the online questionnaire in order to gain insight to their online news-site use habits. A summary is presented in Table 3.4.

Table 3.4. Descriptive statistics of news-site use behaviour.

News-site use frequency	"How often do you visit online news sites?"		
	Gazette	Non-Gazette	Overall
Weekly	3.2% (2)[a]	13.3% (2)	5.2% (4)
2-3 times a week	17.7% (11)	33.3% (5)	20.8% (16)
4-6 times a week	4.8% (3)		3.9% (3)
Daily	48.4% (30)	26.7% (4)	44.2% (34)
2-3 times a day	11.3% (7)	20% (3)	13% (10)
More than 3 times a day	14.5% (9)	6.7% (1)	13% (10)
News-site use duration	"How much time do you spend on online news sites?"		
	Gazette	Non-Gazette	Overall
2-4 minutes	3.2% (2)	6.7% (1)	3.9% (3)
About 5 minutes	9.7% (6)	20.0% (3)	11.7% (9)
About 10 minutes	22.6% (14)	13.3% (2)	20.8% (16)
About 15 minutes	37.1% (23)	46.7% (7)	39.0% (30)
About 30 minutes	14.5% (9)	13.3% (2)	14.3% (11)
More	12.9% (8)		10.4% (8)
Commenting frequency	"Do you comment the news and blog entries on news sites?"		
	Gazette	Non-Gazette	Overall
Never	54.8% (34)	80.0% (12)	59.7% (46)
Incidentally	19.4% (12)	20.0% (3)	19.5% (15)
Sometimes	24.2% (15)		19.5% (15)
Often			
Very frequently	1.6% (1)		1.3% (1)
Online news preference	"Do you prefer online news sites over printed newspapers?" (YES)		
Gazette	Non-Gazette	Overall	
62.9% (39)	40.0% (6)	58.4% (45)	
Regular purchase of printed papers	"Do you purchase printed newspapers regularly or are you subscribed to any daily papers?" (YES)		
Gazette	Non-Gazette	Overall	
41.9% (26)	73.3% (11)	48.1% (37)	

[a] *n*

Fifty-four participants (70%) reported daily or more frequent access to online news sites. Daily access to news sites was the most common answer to the question regarding the frequency of news-site use in the pooled sample (44%, *n* = 34). Forty-nine (64%) reported 15 minutes or more time spent on online news sites. The most-common answer for news-site use duration was *about 15 minutes* in both subsamples (39% in the pooled sample). For news-site use duration, a *1 minute* response category was also available, but no participant indicated less than 2-4 minutes spent on news sites.

Commenting news and blog entries, thus contributing to the site content by sharing thoughts, opinions and discussing topics was a functionality that the majority of participants did not use. Forty-six (60%) reported to never comment on news and blog entries. However, 40% (*n* = 31) generated content to some extent. Fifteen participants (20%) reported to comment incidentally and fifteen to comment sometimes. Only one participant reported to comment very frequently.

Almost two thirds of Gazette readers (63%) reported to prefer online news sites over printed newspapers and 42% reported to be subscribed to daily papers or to purchase printed newspapers regularly. These percentages indicate that Gazette users had a preference towards online news site use and less than half accessed traditional printed media on a daily basis. The five most-visited news sites of Gazette readers are presented in Table 3.5.

Table 3.5. Five most-frequently visited news sites of Gazette users (*n* = 62).

	First	Second	Third	Fourth	Fifth	Total
Gazette	30.6%(19)	30.6% (19)	11.3% (7)[a]	8.1% (5)	1.6% (1)	82.3% (51)
BBC	40.3% (25)	16.1% (10)	16.1% (10)	1.6% (1)	3.2% (2)	77.4% (48)
Sky News	8.1% (5)	9.7% (6)	8.1% (5)	1.6% (1)	0% (0)	27.4% (17)
The Guardian	3.2% (2)	1.6% (1)	6.5% (4)	1.6% (1)	0% (0)	12.9% (8)
The Times	1.6% (1)	4.8% (3)	1.6% (1)	3.2% (2)	0% (0)	11.3% (7)
Northern Echo	1.6% (1)	4.8% (3)	1.6% (1)	1.6% (1)	1.6% (1)	11.3% (7)
Daily Telegraph	3.2% (2)	3.2% (2)	1.6% (1)	1.6% (1)	0% (0)	9.7% (6)
The Sun	0% (0)	1.6% (1)	3.2% (2)	0% (0)	0% (0)	4.8% (3)
Other	8.1% (5)	21.0% (13)	21.0% (13)	17.7% (11)	9.7% (6)	
No answer	3.2% (2)	6.5% (4)	29.0% (18)	62.9% (39)	83.9% (52)	

[a] *n*

The Gazette was reported at some position in the top-five news sites by the majority of Gazette users (82%). The second-most popular news site was BBC (77%), followed by Sky News (27%). Only one local news site other than Gazette was reported in the top five, The Northern Echo, which was reported by 11% of Gazette users. Twenty-five Gazette users (40%) reported BBC and 19 (31%) reported Gazette as the most-visited news site. However, Gazette was reported in second place by exactly as many participants as in first place, whereas BBC was reported by less (16%) in second place. Considering the first three places all together, Gazette and BBC were reported by exactly as many participants (73%, *n* = 45). Response rates dropped dramatically for the fourth and fifth positions of the five most-visited news sites (63% and 84% *no answer*, respectively), which suggests that the majority of Gazette readers used three or fewer online news sites regularly. Overall, the findings indicate that Gazette users' primary source of local online news was Gazette, and they acquired general online news primarily from the BBC news site.

In summary, participants were predominantly middle-aged, employed people and students, mostly with an average to high level of education. They had used the Internet for approximately ten years and most used it on a daily basis for a wide

range of tasks and activities. The majority used online new sites daily. Most Gazette readers used three or fewer online news sites and they used the Gazette and BBC sites most frequently.

3.5. Analysis of the think-aloud data

3.5.1. Extraction of units of thought

Each participant provided approximately 10 minutes of screen-capture video and audio recording. Some participants indicated that they got bored of browsing the Gazette site before 10 minutes had passed. When that happened, the recording was stopped early in order to collect only meaningful material, resulting from spontaneous use of the site. Other participants provided more than 10 minutes of material, because they were not interrupted after 10 minutes if they were still in the midst of exploring the site. One Gazette reader's recording was damaged during the conversion of screen-capture data into AVI format, thus only twenty-four recordings entered the analysis.

The recordings were watched repeatedly in order to identify meaningful units for transcription. As a result, 190 units of thought were identified (based on Riffe et al., 2005) and transcribed. Units of thought were defined as expressions of participants' opinion or judgement about the site (e.g., content, presentation and usability), or expressions of their experience. Units of thought were also collected where participants described how they usually used the site. Because participants were not instructed explicitly to express their opinion about any specific characteristic or functionality of the site and they were not prompted to do so by the researcher during the sessions, the collected units of thought are regarded as genuine and meaningful. Units are regarded as genuine in the sense that the researcher did not bias the

participants' report of their experiences, and regarded as meaningful in the sense that the participants only commented on aspects of their experience that caught their attention and elicited an expression of opinion. In other words, think-aloud participants only talked about things that mattered to them and played a part in their interaction experience.

All transcribed units included time codes and the screen-capture recordings were used to identify the pages from which the units of thought were derived. Next, the participants' identifiers were removed from the units and, instead, a serial number from 1 to 190 was given to each for identification. The resulting list was read repeatedly by the researcher to identify patterns in the units. Five main categories of units of thought were identified: impression, content, layout, information architecture and diversion[9]. The *content* category was further divided into the sub-categories *information content*, *outdated content* and *advertisement*. The *information architecture* category was divided into the sub-categories *links*, *navigation* and *structure*. The *diversion* category was divided into the sub-categories *distraction*, *confusion* and *loading time*. The categories and sub-categories were defined by the researcher and the units were assigned to them. Two independent raters were given the category and sub-category definitions and completed the task of coding all the units using the categorisation scheme, in order to assess the inter-rater reliability of the categories. The instructions given to the independent raters allowed units of thought to be coded in more than one category.

[9] A full description of categories and sub-categories used in the coding is presented in Appendix 3.3.

Coders were instructed to use the *impression* category for units of thought that included first impressions, overall visual impressions and overall evaluative judgements of the site. Overall visual impressions meant statements about the visual appearance of the site in general, while overall evaluative judgements meant an evaluative statement about the site as a whole.

The *content* category was defined for the coders by three sub-categories that were derived from the preliminary reading of the units by the researcher. These sub-categories were: information content, outdated content and advertisement. The sub-category *information content* included units of thought on textual and multimedia content of the site and its articles (or a lack thereof). Coders were instructed to assign every unit of thought regarding the content of the site in general or the content of an individual article in particular to *information content*. Units of thought regarding text, picture and video content all fell into this sub-category. The coders were instructed to use the sub-category *outdated content* for units of thought in which participants identified elements of textual or multimedia content as not being up-to-date. The coders were asked to categorise units of thought regarding advertisements in a distinct sub-category. Nevertheless, advertisements too were regarded as content elements of the site. It was expected that the sub-category *advertisement* would overlap with the sub-category *distraction* of the *diversion* category, because many advertisements were described as distracting in several units.

The *layout* main category included units of thought regarding the layout of text and graphic elements on the site, and article presentation style. Coders were instructed to assign units here that regarded, for example, how the text was broken up into

paragraphs and pages, where certain graphic elements like advertisements and navigation controls were located and how individual pages were presented. The coders judged the *layout* and *information architecture* categories very similar and reported uncertainty when deciding whether a unit fell into one category, the other or both. This uncertainty was also reflected in the inter-rater reliability scores of the two categories, but after collapsing the categories, the reliability scores reached a satisfactory value. The categories *layout* and *information architecture*, however, are presented here separately, because they represent two distinct levels of organisation of graphics and content elements: *layout* concerns individual pages and *information architecture* the entire site. For these two categories, only units on which all three coders agreed upon are presented in this discussion as typical quote examples.

The *information architecture* category was defined by its three sub-categories for the coders. The *links* sub-category included units of thought regarding hyperlinks and connectivity of the site with other Web sites. The *navigation* sub-category included units regarding navigation between individual pages of the site and units regarding the search functionality. Finally, the *structure* sub-category included units of thought regarding information organisation (or lack thereof, e.g., information overload) on the site, as opposed to units regarding the organisation of information (and graphic elements) within individual articles, which were to be assigned to the *layout* category. This distinction between the *layout* category and the *structure* sub-category of the *information architecture* category was later identified as a source of confusion for the coders, because it was not always easy to make a clear decision. The coders were asked to assign any units of thought to the *structure* sub-category that were related to how information was organised or how it should have been organised or

categorised on the site, as well as any units regarding the lack of proper information organisation.

The *diversion* category was defined by its three sub-categories: distraction, confusion and loading time. The coders were instructed to use the *distraction* sub-category for units of thought that suggested a participant being distracted by something, for example, advertisements and flashing graphic elements on the site. The *confusion* sub-category was used in the case of units of thought that suggested confusion in a participant. The *loading time* sub-category was used whenever a participant made a comment about loading time.

3.5.2. Reliability of coding

Various inter-rater reliability coefficients were calculated in order to assess the validity of the categories[10]. Inter-rater reliability coefficients were calculated using *ReCal*, an online inter-coder reliability Web service (*http://dfreelon.org/utils/recalfront/*). *ReCal* is a free-to-use, online application which offers a wide variety of coefficients to assess the degree of agreement between two or more independent coders (Freelon, 2010). The inter-rater coefficients for the researcher and two independent coders are presented in Table 3.6.

[10] Inter-rater reliability coefficients were not calculated for the sub-categories, because the categories were in part defined by their sub-categories, and all units assigned to the sub-categories were also automatically assigned to the corresponding main categories.

Table 3.6. Inter-rater reliability coefficients of the categories with three coders.

	Category				
	Impression	Content	Layout	Information architecture	Diversion
Average pairwise percent agreement	87%	74%	86%	79%	87%
Pairwise pct. agr. 1&3	94%	88%	86%	82%	91%
Pairwise pct. agr. 1&2	86%	68%	85%	76%	86%
Pairwise pct. agr. 2&3	82%	67%	86%	78%	84%
Fleiss' Kappa	0.47	0.47	0.43	0.39	0.58
Observed agreement	0.87	0.74	0.86	0.79	0.87
Expected agreement	0.76	0.52	0.75	0.65	0.68
Average pairwise Cohen's Kappa	0.49	0.46	0.43	0.39	0.59
Pairwise CK 1&3	0.71	0.75	0.41	0.49	0.68
Pairwise CK 1&2	0.45	0.34	0.42	0.31	0.56
Pairwise CK 2&3	0.30	0.31	0.46	0.37	0.52
Krippendorff's Alpha	0.47	0.47	0.43	0.39	0.58

Note. Number of raters = 3; number of cases = 190; number of coding decisions = 570.

Fleiss' Kappa and Cohen's Kappa reliability coefficients indicate the level of agreement among coders with chance agreement taken into account. The general formula of the Fleiss' Kappa and Cohen's Kappa coefficients is identical: the observed amount of agreement between raters above chance ($P_{observed}$ - $P_{expected}$) is divided by the level of agreement attainable above chance (1 - $P_{expected}$). A value of zero indicates no agreement and a value of 1 indicates complete agreement among the raters. Cohen's Kappa measures the agreement between two raters, while Fleiss' Kappa is capable of expressing the level of agreement between more than two raters. The two coefficients differ in how expected agreement is calculated (see Cohen, 1960; Fleiss, 1971). Nevertheless, they provided very similar values for the present data. According to Landis and Koch's (1977) guidelines, Kappa values of between zero and 0.20 are characterised as slight agreement, 0.21-0.40 as fair, 0.41-0.60 as moderate, 0.61-0.80 as substantial and 0.81-1 as almost perfect agreement. As opposed to the Kappa coefficients, Krippendorff's Alpha is based on

a formula expressed in terms of disagreement instead of agreement (Hayes & Krippendorff, 2007).

According to the Kappa values, the categories produced moderate level agreements for all three raters. However, the coding instructions given to the coders allowed for the units of thought to be assigned into any number of categories, not just one. An *unassigned* category was also provided, to be used for units that could not be assigned by a coder to any of the other categories. Therefore, any unit could be potentially assigned to any number of categories between 0 and 5. Furthermore, there was no indication of how many units should be assigned to each of the categories. For each category, a separate data file was produced, with 190 rows representing each unit and 3 columns representing the three coders. A unit of thought received a score of 1 under a coder if the coder assigned that unit into that particular category, and a score of 0 if not. Because of the large amount of 0 codes in the tables, the expected agreement values were inflated, producing lower Kappa coefficients with relatively high levels of percent agreement (see Table 3.6 for expected agreement values for Fleiss' Kappa). For example, for the *impression* category, 87% observed agreement was accompanied by 76% expected agreement. Because there was no indication whatsoever on the number of units to be categorised in any of the categories, completely random filling of the tables with 190 rows would hardly be expected to reach a 76% agreement. On average, the three coders assigned 26 units to the *impression* category (*SD* = 8.19). The large amount of matching 0 values in the tables produced an unreasonably high amount of agreement expected by chance. For this reason, the percent-agreement values are considered to be more accurate gauges of inter-rater reliability in the present case.

From the pairwise percent agreement values of Table 3.6 it is apparent that Coders 1 and 3 had a noticeably higher level of agreement for all the categories, except for *layout*. Coder 2 had lower agreement values with Coders 1 and 3 than the latter two between each other. This also reflects in the pairwise Cohen's Kappa values. Therefore, Coder 3 was identified as having higher levels of disagreement with other coders on most of the categories. The reliability coefficients for Coders 1 and 3 are presented separately in Table 3.7.

Table 3.7. Inter-rater reliability coefficients of the categories with Coders 1 and 3.

	Category				
	Impression	Content	Layout	Information architecture	Diversion
Percent agreement	92%	88%	86%	82%	91%
Scott's Pi	0.71	0.75	0.41	0.49	0.68
Cohen's Kappa	0.71	0.75	0.41	0.49	0.68
Krippendorff's Alpha	0.71	0.75	0.41	0.50	0.68
Number of agreements	179	167	163	155	172
Number of disagreements	11	23	27	35	18

Note. Number of raters = 2; number of cases = 190; number of coding decisions = 380.

With the exclusion of Coder 2 from the reliability analysis, the reliability coefficients of the *impression*, *content* and *diversion* categories rose to a substantial level (above 0.61). The reliability values of the *layout* and *information architecture* categories remained moderate. However, all coders reported that they perceived the *layout* and *information architecture* categories very similar and found it problematic to separate them from one another. Both categories pertain to the organisation and presentation of content and graphic elements; *layout* concerns individual pages, whereas *information architecture* concerns the organisation of elements on the site. Moreover, the reliability coefficients showed to be the lowest for these two

categories, indicating lower levels of agreement between the raters whether a particular unit was a member of the *layout* or the *information architecture* category.

In order to investigate if *layout* and *information architecture* are best treated as one category, the two categories were collapsed and inter-rater reliability scores were calculated. The combined Scott's Pi, Cohen's Kappa and Krippendorff's Alpha values have risen to 0.61, which means a 0.2 rise in values for *layout* and 0.1 for *information architecture*. These results suggest that lower agreement values for the categories *layout* and *information architecture* were in part due to the similarities between them. However, the two categories were kept separate in further analyses for conceptual reasons, but only those quotes were used as typical examples for each category on which both Coders 1 and 3 agreed.

In addition to the inter-rater reliability coefficients, test-retest stability values were also calculated, using two separate category assignments of the units by Coder 1 with two months between the two coding sessions. The results are presented in Table 3.8. The reliability coefficients of all categories were substantial, indicating sufficient stability of the categories over time.

Table 3.8. Test-retest stability coefficients of the categories after two months.

	Category				
	Impression	Content	Layout	Information architecture	Diversion
Percent agreement	96%	90%	91%	91%	95%
Scott's Pi	0.79	0.78	0.66	0.74	0.84
Cohen's Kappa	0.79	0.78	0.66	0.74	0.84
Krippendorff's Alpha	0.79	0.78	0.66	0.74	0.84
Number of agreements	182	170	172	172	181
Number of disagreements	8	20	18	18	9

Note. Number of cases = 190; number of coding decisions = 380.

3.5.3. Size and prevalence of the categories

On average, the three coders assigned 92% (*M* = 175; *SD* = 2.65) of all 190 units of thought to at least one of the categories. The marginal use of the *unassigned* category by the coders indicates that the categories accounted for the data well and the majority of units of thought were categorised using the coding scheme. The *size* of the categories was expressed by the average number and percentage of units of thought assigned to them (averaged over coders). The *prevalence* of the categories was defined as the number and percentage of participants that had a unit of thought assigned to each of the categories. The size and prevalence of the categories based on the ratings of all three coders are presented in Table 3.9.

Table 3.9. Size and prevalence of the categories.

	Category									
	Impression		Content		Layout		Information architecture		Diversion	
Mean *N* of units of thought (SD)	26.00 (8.19)		75.33 (2.52)		28.00 (3.00)		43.33 (4.04)		37.67 (6.66)	
Mean % of all units of thought (SD)	13.68% (4.31)		39.65% (1.32)		14.74% (1.58)		22.81% (2.13)		19.83% (3.50)	
N of units agreed upon by coders	16		65		13		27		25	
N of participants reporting to category	11 (46%)		19 (79%)		8 (33%)		14 (58%)		11 (46%)	
	N[a]	G[b]	N	G	N	G	N	G	N	G
N of units per group	14	2	43	22	9	4	20	7	21	4
N of participants reporting to category	9	2	12	7	5	3	10	4	9	2
Odds ratio	5.24		1.14		1.00		2.50		5.24	

[a]Non-Gazette reader

[b]Gazette reader

Content proved to be the largest of the categories with 40% of all units assigned to it. The majority of participants' verbalisations referred to the material presented on the site. The category *content* also showed to be the most prevalent one, containing at

least one unit of thought from 79% of the participants. *Information architecture* was the second largest category with 23% of the units of thought assigned to it, followed by *diversion* (20%), *layout* (15%) and *impression* (14%). *Information architecture* was also the second-most prevalent category, containing at least one unit of thought from 60% of the participants. The *impression* and *diversion* categories were tied for the third-most prevalent category with 46% and the least prevalent category was *layout* with 29%. In the following sections, quote examples are presented for the categories and their sub-categories from the transcribed units of thought.

The numbers of units of thought that were agreed upon by Coders 1 and 3 to be members of the categories are also presented separately for non-Gazette readers (i.e., novice users) and Gazette readers (i.e., expert users) in Table 3.9. Because the expected frequencies were lower than five for at least 25% of the cells for each category, χ^2 tests or loglinear analysis could not be conducted to test the differences between the numbers of people reporting to each category from the two groups (Field, 2009). However, the differences in the number of units left by Gazette and non-Gazette readers in the categories are expressed as odds ratios in Table 3.9, which can be interpreted as the number of times the odds of a novice reader leaving a unit in a category are greater than the odds of an expert reader[11]. For the *impression* and *diversion* categories, the odds of non-Gazette readers providing a unit of thought were 5.24 times higher than Gazette readers. Similarly, novice users reported more units to the category *information architecture* than expert users. There were no notable difference in the number of units reported by novice and expert

[11] Note that the sample sizes were not equal for the two groups. Units of thought were collected from 15 non-Gazette readers and 9 Gazette readers.

users to the categories *content* and *layout*. All units of thought assigned to the categories are presented in Appendix 3.4. To provide an overview, a summary of categories and sub-categories of units of thought is presented in Table 3.10. Quotes from the transcribed protocols to each category and sub-category are presented in the following sections.

Table 3.10. A summary of categories of units of thought.

Category	Sub-category	Short description of units' content
Impression	N/A	First impressions, overall visual impressions and overall evaluative judgements
Content	Information content	Site and article content (text and multimedia), or lack thereof
	Outdated content	Comments on content being out of date
	Advertisement	Comments on advertisements
Layout	N/A	Layout of text and graphic elements on the site and article presentation style
Information architecture	Links	Hyperlinks and connectivity of the site with other Web sites
	Navigation	Navigation between pages of the site and comments on the search functionality
	Structure	Information organisation on the site (not within individual articles), or lack thereof (e.g., information overload)
Diversion	Distraction	Comments suggesting being distracted (e.g., by irrelevant content and graphics elements)
	Confusion	Comments suggesting confusion (e.g., by page layout)
	Loading time	Comments on loading time

3.5.4. Impression

According to all three coders, on average 14% of the units of thought were assigned to the *impression* category. These units were produced by 11 of the 24 participants (46%) who provided data for the analysis. The *impression* category is the smallest and one-but-the-least prevalent of all the categories. All three coders agreed unequivocally on 9 units of thought to be a member of this category, while Coders 1 and 3 agreed on 16 units. On average, the three coders assigned 26 units of thought to this category (*SD* = 8.19).

Although *impression* was the smallest category, almost half of the users spontaneously expressed at least one overall evaluative statement about the site. Four of these statements regarded the visual or aesthetic appearance of the site. An example of a quote: “It’s a bright site” (Participant 2). Another example: “At least it’s very colourful, not like the other news once I’ve been on” (Participant 25). The amount of information presented to the users on the screen also elicited several units of thought, for example, “Lots of information here... yeah, the information here is lots. Which I like about Web sites” (Participant 8). Another example of a quote: “Lots of things are going on” (Participant 13). The following example also contains an element of confusion: “My first reaction’s there’s an awful lot of information and I’m really not sure where I’d go first” (Participant 12). Six of the units contained outright evaluative judgements or attitudes about the site, for example: “This is a nice site. It’s not as bad as it used to be four years ago, when I think it was the last time I visited the Gazette site” (Participant 24). Another example, with an even greater degree of evaluation: “I don’t like this Web site at all” (Participant 25).

Out of the 16 units agreed upon by Coders 1 and 3 in this category, 14 were made on the home page of the site and 12 were made within the first minute of the browsing sessions. The protocol was divided into one-minute segments and the number of units was counted for each category to examine possible trends in reporting different aspects of experience. The number of units in each category is presented in Figure 3.1. The findings suggest that overall impressions form relatively early during use and may not be based on a thorough exploration of the site. Nearly half of the participants provided at least one unit of thought in the *impression* category, which indicates the plausibility of including measures of overall evaluative judgements in further model development.

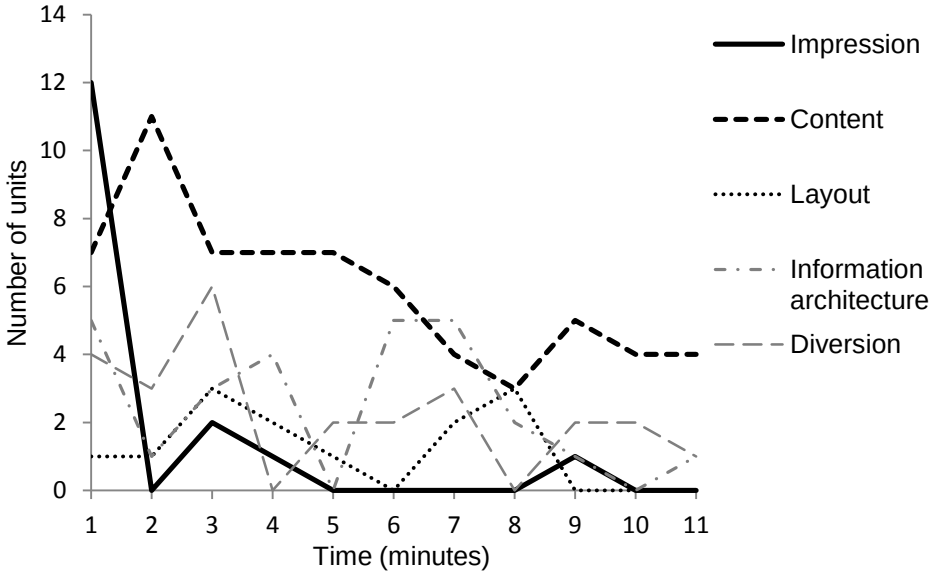


Figure 3.1. Number of units of thought reported by participants to each category from the first to the eleventh minute.

3.5.5. Content

According to all three coders, on average 40% of the units of thought were assigned to the *content* category. These units were produced by 19 of the 24 participants (79%). All three coders agreed unequivocally on 39 units of thought to be a member of this category, while Coders 1 and 3 agreed on 65 units. On average, all three coders assigned 75.33 units to the *content* category (*SD* = 2.52). The category *content* was the largest and most prevalent of all categories. The number of units and the high prevalence of the category indicate that the majority of participants were predominantly concerned with the content of the site during their interaction, which suggests that site content was the most important factor of quality of interaction with the news site.

The following example of a quote illustrates an expression of opinion of the overall content of the site: “It’s all antisocial behaviour or drugs” (Participant 2). Another example of an expression of opinion on information content, but this time regarding a picture gallery rather than the entire site: “This is quite interesting... local interest, history, I find it quite fascinating” (Participant 7). Finally, a unit of thought on the sports section of the site: “Kind of good to see really that it doesn’t... football just doesn’t dominate it” (Participant 20). The collected units were abundant with similar comments on information and multimedia content of the entire site, different sections and individual articles.

There were eight units of thought in which participants identified a piece of content to be out of date, as illustrated by the following quotes. “But what you notice also on this is quite often the so-called latest news is from the day before” (Participant 3). “So you could ask is this information still... still relevant” (Participant 8). “The latest

image gallery is March 2009, which doesn't appear very latest to me" (Participant 16). These units of thought were assigned to the sub-category *outdated content*. The participants were sensitive to the relevance and timeliness of the information presented to them.

A high number of units of thought (M = 18.67; SD = 3.21) were collected about advertisements as content elements of the site. The Gazette Live site used in the study had many advertisements and followed the business model of providing free content and generating revenue by selling advertisement space. An example of a quote about advertisements: "The first thing I notice is the army advert again. The adverts are very prominent on the top, it's a good place to advertise, looks like. And a little film there ..." (Participant 13).

The following quotes illustrate the case when advertisements are also identified as distracters. "That advert for Lidl keeps distracting me on the right hand side; I keep wanting to look at that" (Participant 6). "The advertising is really annoying. It's actually quite big on the screen and so ... I'm not used to that ... I'm used to two smaller sidebars really, but that's quite distracting. Oh, I'll just ignore it" (Participant 16). Finally, the following example of a quote illustrates a case where a piece of advertisement elicited a negative remark with its content. "Again, more adverts. [...] It's kind of the first you see when you look at it. There's adverts about giving in your gold for like, money... pawnbrokers. It seems like they just preying on people 'cause they know they've got less money now because there is a recession" (Participant 10).

The large number of units of thought regarding advertisements indicates that these elements played a significant part in users' interaction with the site, predominantly as distracters. Generally, units about advertisements had a negative valence (valence

analysis of units of thought is presented in Section 3.5.9). Advertisements, when characterised as distracters, can be regarded as irrelevant content elements that may decrease the user-perceived quality of interaction. Therefore, news sites that follow another business model (e.g., a subscription-based model or government funding) and do not include separate advertisement spaces may be regarded as containing less irrelevant content. In summary, the size, prevalence and the sub-categories of the *content* category suggest the importance of the quality, timeliness and relevance of content for interaction experience with news sites.

3.5.6. Layout

According to all three coders, on average 15% of the units of thought were assigned to the *layout* category. These units were produced by seven of the 24 participants (29%). All three Coders agreed unequivocally on 9 units of thought to be a member of this category, while Coders 1 and 3 agreed on 13 units. On average, all three coders assigned 28 units of thought to the *layout* category (*SD* = 3.00).

The following example of a quote illustrates a relatively frequently reoccurring layout issue, the presentation of short articles on several pages. “Okay, so the news story is split over two pages. That’s a bit irritating. If there is... there’s a lot of space at the bottom for... oh, there’s loads of space at the bottom. [...] I think it’d be much better if that was just over one page, if the story was just only one place and you could just read all of it, rather than having to flip to another page” (Participant 15). Another example of layout-related issues: “...but again, I find it there’s still a lot of information flashing in front of me, which I don’t particularly like. And the writing is really quite small for me anyway” (Participant 12). This participant was distracted by irrelevant graphic elements on the page from the main content (an article in this case), which

was presented in a less apparent manner than the accompanying material. These examples illustrate that layout issues can frustrate users and divert users' attention from browsing site content, decreasing the user-perceived quality of interaction.

Finally, a positive unit of thought on the consistent design of the site: "Also what is very consistent is the Gazette... they use the red and the blue and it does seem to flow throughout the sequence. All the quick links are in red and all the major headlines are in blue and the text seems to be in blue and... So you have, there's a consistency of that" (Participant 22). This unit of thought illustrates that good layout and consistent design may not go unnoticed and can elicit positive remarks from users.

In summary, the *layout* category consisted of units of thought on graphics and presentation design of the interface. Almost 30% of the participants provided a unit of thought to this category, indicating that page layout and presentation attributes caught the attention of a significant proportion of users. Units with both positive and negative valence were collected. Issues related to layout may act as distracters and may frustrate users. Good layout presumably facilitates the quality of interaction predominantly through allowing for the unhindered exploration of content.

3.5.7. Information architecture

According to all three coders, on average 23% of units of thought were assigned to the *information architecture* category. These units were produced by 14 of the 24 participants (58%). All three coders agreed unequivocally on 14 units of thought to be a member of this category, while Coders 1 and 3 agreed on 27 units. On average, all three coders assigned 43.33 units of thought to the *information*

architecture category (*SD* = 4.04). *Information architecture* was the second largest and most prevalent category after *content*.

Four quotes follow as typical examples of the sub-category *links*. "They've got official Web site, Facebook, a blog, then I get MySpace or Twitter of flickr... Quite a few different connecting to quite popular social network sites, that's good" (Participant 4). This quote contains a positive remark of the connectivity of the site with other, popular social network sites. The following quote suggests confusion in another participant when faced with a wide assortment of links and navigational choices: "...especially when there is so much information on the Web site, on the home page and the various links... I just can't see which... which is which" (Participant 12). The next quote points out the usefulness of links leading to other sites from a particular page of the news site: "Ah, then there's links to pages outside as well [points to middle bar], which is... that's good. It's useful" (Participant 15). Finally, in the last example a participant is looking for related links and encounters a high number of choices presented in a style that makes it hard to pick out the piece of information required: "Looking at the bottom of the homepage to see if there's anything about sports pages... Yes, I've found one but it's very, very small and there's an awful lot of links and hyperlinks here" (Participant 12).

The *navigation* sub-category was strongly related to the *links* sub-category in the content of units of thought. Indeed, many of the units of one sub-category were assigned to the other and vice versa. For this reason, the sub-categories are not strictly separated from each other, but rather used to describe the categories in closer detail. The following quote from the *navigation* sub-category illustrates a participant's perception of ease of navigation: "So probably I'll go back to... I go back

to UK and world news, and that's at the top at one of the tabs which is quite easy to navigate around ..." (Participant 11). A common navigational aid on news sites is the inclusion of a search text-box. Often a popular search engine is embedded to serve this function. The inclusion of Google search on Gazette Live received positive comments from the participants. An example of a quote: "Oh, you've got a Google search here, that's a site which I constantly use. I'm glad that they've got that on here" (Participant 8).

The following quote illustrates the *structure* sub-category by highlighting a positive experience related to information organisation: "So this must be the main menu for all the football news. That's a good idea, I like having that separate to the sports, so if you're just interested in Middlesbrough Football Club" (Participant 7). An important issue regarding this sub-category that influenced experience of use was a lack of proper organisation, which might result in information overload and consequently in feeling 'lost' and frustrated. The following quote illustrates this matter: "I find the Web site a bit heavy as well. It's too much information going on and that makes me a bit irritated" (Participant 24). The effects of lack of organisation may be in part compensated by providing a search function, but the ease and experience of browsing may still be compromised.

In summary, the size and prevalence of the *information architecture* category indicates the importance of content organisation, providing a manageable number of links related to the content of the site and providing search facilities to allow for quick access to desired pieces of information, bypassing the navigation structure of the site. Typical issues with information architecture included feeling lost, confused and frustrated. Positive units of thought on information architecture were frequently

concerned with the connectivity of the site with other sites, good organisation of site content and ease of navigation.

3.5.8. Diversion

According to all three coders, on average 20% of the units of thought were assigned to the *diversion* category. These units were produced by 11 of the 24 participants (46%). All three coders agreed on 20 units of thought to be a member of this category, while Coders 1 and 3 agreed on 25 units. On average, all three coders assigned 37.67 units of thought to the *diversion* category (*SD* = 6.66).

As discussed before in relation to the *advertisement* sub-category of the *content* category, distraction was most commonly associated with advertisements on the site. An example of a quote: “Just linking on to sports, there’s an awful lot of information here. There’s a lot of... flash and links at the top... about the army. There’s a lot of other links down the side that’s quite off-putting, actually” (Participant 12). The presentation of a high number of other content elements and links was also found to be distracting. The following quote illustrates how such layout-related issues may act as distracters: “There is far too much information going on at once. Which is good, ‘cause it kind of makes me want to read everything, but it’s taking my mind off what I’m actually looking for” (Participant 6).

Quotes related to the *confusion* sub-category were most commonly associated with issues regarding the layout of individual pages and the information architecture of the site, as illustrated by the following quotes. “I’ve just noticed that the date has now changed on these headlines I’m reading. It did change; some of them were on the 14 of July. Which has confused me a little bit, as I thought it was all in chronological... well, going back in a chronological order” (Participant 11). “Oh.

Right. There's a local community Web site with lots of initial post codes and letters and numbers... which really looks quite confusing" (Participant 12). "...but there's so much information I'm really finding it quite hard to see if there's any links anywhere else about market times ..." (Participant 12).

Finally, eight units of thoughts concerned the loading time of pages. Five of the units were negative, for example: "It takes too long to load" (Participant 25). All participants used the same computer during the think-aloud session with high-speed Internet connection, which was noted by some participants, for example: "Not taking too long to load" (Participant 4). This example illustrates the relevance of page-loading time to participants' interaction experience.

In summary, units of thought in the *diversion* category are associated with issues regarding the categories of *content*, *layout* and *information architecture*, except for the units of thought in the *loading time* sub-category, which shows no overlap with any other type of unit of thought. Therefore, distraction and confusion are presumably related to cases when the site does not meet a particular participant's preferences regarding the content, presentation style and information organisation of the site. These issues need not to be addressed separately from the above mentioned categories in further psychometric measurement, because they merely represent the negative cases of the categories. The measurement of users' perception of loading time, however, needs to be measured separately to assess its role in interaction experience with news sites.

3.5.9. Valence of units of thought

Valence analysis was carried out on the 146 units of thought that were agreed upon by the coders (see Table 3.11). Of all units, 44 were rated as positive, 79 as

negative and 23 as neutral. The overall ratio of negative units to positive was 1.80. However, valence of units was not evenly distributed among categories. The category that received the most negative units of thought was *diversion*, where negative units outnumbered positive ones four to one. Timeline analysis of units by valence in Figure 3.2 shows that positive and negative units followed approximately similar distribution; however, there was a drop in the number of negative units after the fourth minute. There was a slight build-up of units with negative valence in the first four minutes, whereas the number of units with positive valence dropped after the first minute and followed a more even distribution.

Table 3.11. The number of units of thought in each category by valence.

Category	Positive	Negative	Neutral	Ratio (-/+)
Impression	4	9	3	2.25
Content	19	30	16	1.59
Layout	5	5	3	1.00
Inf. architecture	11	15	1	1.36
Diversion	5	20	0	4.00

Units of thought with positive and negative valence were slightly unevenly distributed among users and non-users of Gazette. There were 14 positive and 16 negative units from Gazette users, whereas there were 30 positive and 63 negative units from non-Gazette users. There was no significant association between user-group (Gazette or non-Gazette) and valence (positive or negative) $\chi^2(1) = 2.05$, $p = .15$; however, based on the examination of odds ratios, the odds of Gazette readers providing a positive unit of thought were 1.84 times higher than non-Gazette readers, and the odds of non-Gazette readers providing a negative unit were 1.14 times higher than Gazette readers. These findings suggest that the two groups were

nearly as likely to express a negative comment about the site, but regular users were more likely to express positive comments too.

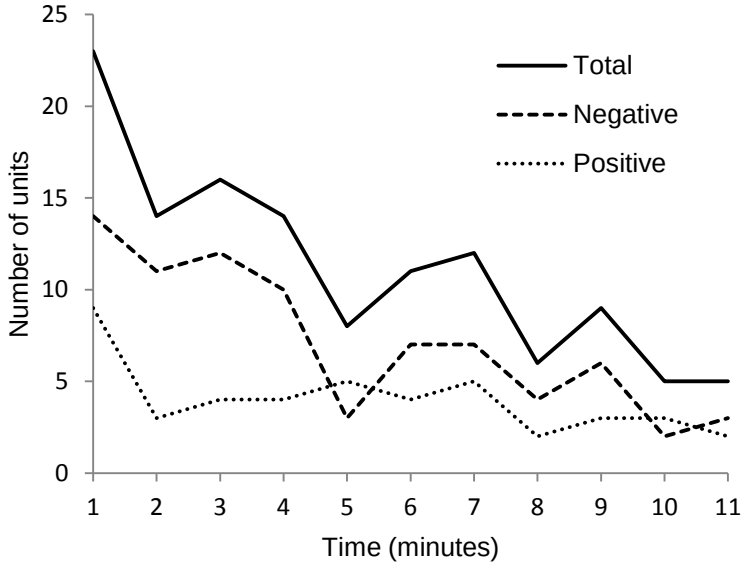


Figure 3.2. Number of units of thought reported by both groups in all categories from the first to the eleventh minute.

3.6. Analysis of post-use questionnaires

3.6.1. Internal consistency of scales and between-group differences

Following the think-aloud task, the participants were asked to complete four interaction-experience questionnaires regarding their use of the site: *perceived aesthetics* to measure perceived artefact characteristics; Attrakdiff2 to assess the participants' perception of the site's *pragmatic quality* and *hedonic quality*, and the evaluation of *beauty* and *goodness*; *perceived enjoyment* and *intensity of flow* to assess the participants' ratings of their quality of interaction. Note that the following analyses are based on 25 participants (10 regular users of Gazette and 15 non-Gazette users) and they are reported for exploratory purposes and to guide the further research phases of the project reported in this thesis. All interaction

experience scales showed high reliability (Cronbach's alpha ranging from .79 to .96). The reliability coefficient of each scale is presented in Table 3.12.

Table 3.12. Internal consistency of interaction experience measures.

Scale	*Cronbach's alpha*
Perceived enjoyment	.96
Intensity of flow	.84
Perceived aesthetics	
Classical aesthetics	.90
Expressive aesthetics	.79
AttrakDiff2	
Pragmatic quality	.85
Hedonic quality	.86

The differences between the answers of the two groups were investigated using unrelated *t* tests on each measure. The tests revealed several significant between-group differences. The results of the independent samples *t* tests are summarised in Table 3.13.

Table 3.13. Independent samples *t* test results of interaction experience questionnaires.

Variable	Gazette Mean (SE)	Non-Gazette Mean (SE)	Mean difference	*t* (23)	*p*	Effect size (*r*)
Perceived enjoyment	4.70 (.47)	2.44 (.36)	2.26	3.89	.001	.63
Intensity of Flow	4.98 (.21)	4.19 (.33)	0.79	1.78	.089	.35
Classical aesthetics	4.36 (.34)	2.96 (.37)	1.40	2.64	.015	.48
Expressive aesthetics	3.68 (.26)	2.77 (.27)	0.91	2.34	.029	.44
Pragmatic quality	4.85 (.25)	4.03 (.41)	0.82	1.50	.147	.30
Hedonic quality	4.15 (.26)	3.20 (.27)	0.95	2.43	.024	.45
Beauty	4.00 (.21)	3.20 (.31)	0.80	2.13	.045	.41
Goodness	4.80 (.33)	3.27 (.44)	1.53	2.53	.019	.47

Note. All variables were measured on a 7-point scale.

Regular readers of Gazette experienced a higher level of *perceived enjoyment* (*M* = 4.70, *SE* = 0.47) than non-Gazette readers (*M* = 2.44, *SE* = 0.36). The difference between the two groups was significant, $t(23) = -3.887$, $p < 0.001$, and represented a

large effect size[12], *r* = .63. Regular Gazette users had also experienced a higher level of *intensity of flow* (*M* = 4.98, *SE* = 0.21) than non-Gazette readers (*M* = 4.19, *SE* = 0.33). This difference was not significant, *t* (23) = -1.776, *p* = .09; however, it did represent a medium effect size, *r* = .35.

The perception of Web-site aesthetics also differed significantly between the two groups. Regular readers of Gazette rated the site higher on the *classical aesthetics* dimension (*M* = 4.36, *SE* = 0.34) than non-Gazette readers (*M* = 2.96, *SE* = 0.37). This difference was significant, *t* (23) = -2.638, *p* = .02, and represented a medium effect size, *r* = .48. Regular readers of Gazette also rated the Gazette site higher on the *expressive aesthetics* dimension (*M* = 3.68, *SE* = 0.26) than non-Gazette readers (*M* = 2.77, *SE* = 0.27). This difference was significant, *t* (23) = -2.335, *p* = .03, and represented a medium effect size, *r* = .44.

Pragmatic quality ratings of Gazette readers and non-Gazette readers did not differ significantly, *t* (23) = -1.503, *p* = .15. However, the Gazette site received higher *hedonic quality* ratings from Gazette readers (*M* = 4.15, *SE* = 0.26) than from non-Gazette readers (*M* = 3.20, *SE* = 0.27). This difference was significant, *t* (23) = -2.425, *p* = .02, and indicated a medium effect size, *r* = .45. With regards to overall quality judgements, regular readers of Gazette rated the Gazette site more *beautiful* (*M* = 4.00, *SE* = 0.21) than non-Gazette readers (*M* = 3.20, *SE* = 0.31). This difference was significant, *t* (23) = -2.126, *p* = .05, and represented a medium effect size, *r* = .41. The Gazette site received higher levels of *goodness* rating form Gazette readers (*M* = 4.80, *SE* = 0.33) than from non-Gazette readers (*M* = 3.27,

[12] Cohen's (1988) conventions for effect size were used: 0.1 for small, 0.3 for medium and 0.5 for large.

SE = 0.44). This difference was significant, t (23) = -2.534; p = .019, and represented a medium effect size, r = .47.

The between-group differences indicate that, in general, regular users had a higher-quality experience than non-users. Regular Gazette readers reported to have experienced significantly higher levels of *enjoyment* and rated the site higher on *goodness* and *beauty*. Regular users of Gazette also rated the site higher on the *hedonic quality* scale, while the two groups did not differ significantly in their *pragmatic quality* ratings. The Gazette site received higher ratings on both subscales of *perceived aesthetics* from its regular users. These findings indicate the importance of taking into account the participants' level of experience and familiarity with a particular artefact; otherwise, a high variability in responses to interaction experience scales can be expected due to individual differences in use habits. However, the groups of the present study represent two extremes: non-users and regular users of a particular artefact. Nevertheless, it seems plausible that more frequent use of an artefact is accompanied by higher quality of interaction experience, especially when the use of the artefact is voluntary. It may be that high-quality interaction experience inspires more frequent use, or the increasing familiarity of a particular artefact with more frequent use enhances the quality of interaction. The relationship between frequency of use and interaction experience and their relation with need fulfilment (Hassenzahl et al., 2010) are investigated in Study 2 (see Chapter 5).

3.6.2. Correlation and regression analyses of post-use questionnaires

In order to explore the relationships between the measures of interaction experience, correlation coefficients were calculated between the scales. All scales were

significantly correlated (Pearson's correlation coefficients ranging from .44 to .86). The measures of *perceived enjoyment* (PE) and *intensity of flow* (Flow) were used to gauge the quality of interaction. *Pragmatic quality* (PQ) and *hedonic quality* (HQ) are measures of perceived product character[13]. *Classical aesthetics* (AC) and *expressive aesthetics* (AE), on the other hand, are measures of perceived product characteristics or product features. Therefore, perceived product character represents a more general assessment of the system than the two dimensions of *perceived aesthetics*, which only represent a judgement of aesthetic aspects of the site. Finally, *beauty* and *goodness* (Good) represent overall quality judgements of the site. The correlation coefficients are presented in Table 3.14.

Table 3.14. Correlations between interaction experience scales.

	Flow	PE	Beauty	Good	PQ	HQ	AC
PE	.53**						
Beauty	.44**	.66**					
Good	.53**	.80**	.74**				
PQ	.79**	.50 *	.48 *	.64**			
HQ	.43 *	.84**	.72**	.89**	.48 *		
AC	.67**	.80**	.66**	.86**	.83**	.79**	
AE	.45 *	.70**	.55**	.71**	.45 *	.78**	.67**

*p (two-tailed) < .05. **p (two-tailed) < .01.

Intensity of flow showed the highest correlation with *pragmatic quality* (r = .79; p < .01), whereas *perceived enjoyment*, the other measure of interaction quality, showed the highest correlation with *hedonic quality* (r = .84; p < .01). *Intensity of flow* and *perceived enjoyment* were strongly correlated (r = .53; p < .01). The overall quality judgements of beauty and goodness had also showed high correlation (r = .74; p <

[13] Pragmatic quality refers to the system's ability to support the achievement of behavioural goals. Hedonic quality refers to the system's ability to make the interaction pleasurable by fulfilling the needs of users (e.g., stimulation and popularity) (Hassenzahl, 2003).

.01). Both *beauty* and *goodness* were most strongly correlated with *hedonic quality* of the two perceived product character subscales (r = .72 and r = .89, respectively; p < .01). *Classical aesthetics* showed the highest correlation with *pragmatic quality* (r = .83; p < .01), while *expressive aesthetics* was more strongly correlated with *hedonic quality* (r = .78; p < .01). The two *perceived aesthetics* subscales were also strongly correlated with each other (r = .67; p < .01). A series of regression analyses was conducted to further examine the relationships between the interaction experience measures.

Because aesthetic quality as an artefact characteristic has a positive effect on hedonic quality (e.g., Hassenzahl, 2004), the two perceived-aesthetics dimensions were regressed onto *hedonic quality* (see Table 3.15). The regression showed a good fit (R^2_{adj} = .71) and the overall relationship was significant, F (2, 22) = 30.44, p < .001.

Table 3.15. Multiple regression analysis of hedonic quality (DV), classical aesthetics and expressive aesthetics.

Variables	Hedonic (DV)	Classical	Expressive	*b*	β	sr^2 (unique)
Classical	.79			**.35	.49	.13
Expressive	.78	.67		**.46	.45	.11
			Constant =	.91		
Means	3.58	3.52	3.13			R^2 = .74[a]
SD	1.05	1.45	1.04		Adjusted	R^2 = .71
						**R = .86

[a]Unique variability = .24; shared variability = .49

***p* < .01

Given the theoretically justified relationship between the two constructs (van Schaik & Ling, 2011) and because *classical aesthetics* showed the highest correlation with *pragmatic quality*, the relationship between the two variables was investigated using simple regression with *pragmatic quality* as dependent variable. The regression

showed a good fit (R^2_{adj} = .67) and the relationship was significant, F (1, 23) = 49.75, p < .001.

In agreement with Hassenzahl (2004), *hedonic quality* and *pragmatic quality* were regressed onto *goodness* (see Table 3.16). The regression showed a very good fit (R^2_{adj} = .84) and the overall relationship was significant, F (2, 22) = 64.58, p < .001. Although *perceived enjoyment* showed high correlation with *goodness*, its inclusion in the regression did not result in a significant improvement of the model, due to its large amount of shared variance with *hedonic quality*. *Intensity of flow* did not contribute significantly to the regression model either.

Table 3.16. Multiple regression analysis of goodness (DV), pragmatic quality and hedonic quality.

Variables	Goodness (DV)	PQ	HQ	*b*	β	sr^2 (unique)
PQ	.64			**.33	.28	.06
HQ	.89	.48		**1.19	.76	.45
			Constant =	-1.81		
Means	3.88	4.36	3.58			R^2 = .85[a]
SD	1.64	1.37	1.05		Adjusted	R^2 = .84
						**R = .92

[a]Unique variability = .51; shared variability = .35

**p < .01

Beauty was most strongly correlated with *hedonic quality*. The simple regression of *hedonic quality* on *beauty* resulted in a good fit (R^2_{adj} = .50) and a significant relationship, F (1, 23) = 24.65, p < .001. The inclusion of *pragmatic quality* and *intensity of flow* did not contribute significantly to the regression model. A summary of correlation and regression analyses are presented in Figure 3.3.

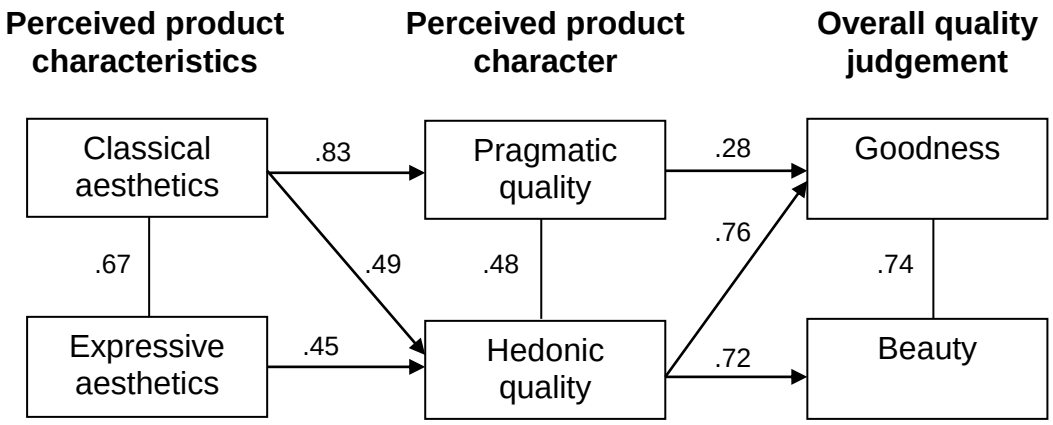


Figure 3.3. Summary of findings of correlation and regression analyses.

Note. Numbers above the arrows are β-weights; numbers next to the vertical lines indicate correlation coefficients.

In summary, the findings of the correlation and regression analyses indicate strong relationships between the different measures of interaction experience. Overall quality judgements could be predicted from the dimensions of perceived product character, which were predicted from the two dimensions of perceived product characteristics included in this study. The overall *goodness* and *beauty* judgements of the site were strongly correlated with measures of quality of interaction, particularly with *perceived enjoyment*. Overall, the results suggest that the inclusion of additional measures, for example, for perceived product characteristics, and collecting data from a larger number of participants in further research phases would provide a useful basis for formulating a model of interaction experience with news sites.

3.7. Discussion

The aim of the exploratory study was to complement the literature review and explore how users of a particular news site interact with the site and report their experience. An online questionnaire was used to recruit participants and to collect data regarding demographics, Internet-use behaviour and news-site-use behaviour of users of the Gazette Live news site. Fifteen non-Gazette users who were regular readers of online news sites have also completed the online questionnaire. Therefore, all participants were regular users of news sites.

Demographic results showed that the participants were experienced users of the Internet with an average of over 10 years of use and most commonly with daily access to the Web. One third of the overall sample indicated a news site as their most-frequently visited Web site and four fifths had indicated at least one news site among their five most-visited Web sites. Most participants accessed news sites on a daily frequency and spent approximately 15 minutes on news sites per visit. These results show that using news sites was an important part of participants' use of the Internet. Moreover, the majority of the full sample preferred online news sites over printed newspapers and less than half reported to purchase any daily papers. Overall, the results of the analysis of the online questionnaire suggest that experience with news sites constitutes a significant portion of regular news-site users' daily online activity and news sites serve as an important source of information to them.

A think-aloud study was conducted with regular users and non-users of the Gazette site in order to explore participants' spontaneous verbal accounts regarding their experience of use. Because the participants browsed the site freely with no set tasks

and they were not prompted to report on any particular aspect of the site or their experience, their verbal comments were regarded as representing genuinely meaningful and undistorted accounts of their experience. The collected units of thought were used to identify aspects of the participants' experience that need to be addressed in a measurement model of interaction experience with online news. Following the think-aloud task, the participants completed a set of interaction-experience questionnaires, selected from the human-computer interaction literature. The analysis of these questionnaires is used to aid the development of a model of interaction experience in the next chapter.

The analysis of the think-aloud recordings yielded five categories: impression, content, layout, information architecture and diversion. The *impression* category contained units of thought about first impressions, visual impressions and overall evaluative judgements of the site. The high prevalence of this category indicates the plausibility of including measures of aesthetics quality and overall evaluations in a measurement model of interaction experience. The category *impression* relates to quick formation and subsequent stability of first impression claimed by Lindgaard et al. (2006) and Tractinsky et al. (2006), but was apparently debunked by van Schaik and Ling (2009) who found a lack of stability and a dependence of stability on context (mode of use). Lindgaard et al. and Tractinsky et al. conceptualised first impression in terms of aesthetics (attractiveness of immediate judgement after 50 or 500 milliseconds or perceptions of classical and expressive aesthetics after a few seconds' viewing). In the current study, the meaning of some units of thought in the impression category seemed to be related to aesthetics, but others to information content of the site, and still others were more general.

Content proved to be the largest and most prevalent category with *information content*, *outdated content* and *advertisement* sub-categories. The size and prevalence of this category indicates that the majority of participants were predominantly concerned with content during their use of the site, which provides grounding for the inclusion of measures that address the quality, relevance and timeliness of information presented on news sites in a measurement model. The categories of *content* and *layout* concern the quality and presentation of content, and correspond to *perceived core service quality*, a direct antecedent of perceived usefulness in the adaptation of the technology acceptance model for news Web sites (Chen & Corkindale, 2008). Because the primary functionality of news sites is to provide news content and *content* was the largest and most prevalent category, indicating that it played a major role in the participants' experience, measures of user-perceived quality of information should be included in the comprehensive psychometric measurement of interaction experience with news sites.

The *layout* category included units of thought regarding the layout of text and graphic elements on the site and article presentation style. The analysis of this category concluded that issues with page layout may serve as distracters and may frustrate users, whereas good layout with consistency in navigation and presentation style presumably facilitates the quality of interaction (Petrie et al., 2009). *Layout* is also strongly connected to the aesthetic appearance of a site. *Layout* is conceptually related to *perceived user-interface design*, which is an important determinant of continued use intention in an adaptation of the technology acceptance model to Web technologies (Cho et al., 2009). *Perceived user-interface design* is recognised as a major factor in Web-technology adoption, because the functionality of a computer system is reflected by its user-interface (Jaspers et al., 2004). Visual aesthetics is

also strongly related to the *layout* category and it plays an important role in interaction-experience models (e.g., Hassenzahl, 2004; van Schaik & Ling, 2009; Zhou & Fu, 2007). Research suggests that the visual aesthetics of computer interfaces is a strong determinant of users' pleasure and satisfaction (Lavie & Tractinsky, 2004). Therefore, measures of perceived interface design and aesthetics are proposed to be included in a measurement model of interaction experience.

The *information architecture* category strongly overlapped with the *layout* category. Furthermore, the *links*, *navigation* and *structure* sub-categories of *information architecture* were also found to be strongly inter-related. *Information architecture* was only second to the *content* category in terms of the number of units of thought it contained, emphasising the importance of content organisation, the quality and appropriate number of links leading to related content elements and other sites, and the search functionality as a navigational aid. Negative units of thought regarding information architecture were connected with confusion, feeling lost and frustrated. The three sub-categories of the *information architecture* category provided evidence for people's experience of three corresponding components of information architecture that have been identified in previous work: labelling system, organisation system and navigation system (Rosenfeld & Morville, 2006). The *information architecture* category summarised the issues of navigating the complex structure of linked content on a news site. According to McDonald and Stevenson (1998), disorientation, the tendency to lose the sense of location on a Web site, is one of the most common problems faced by users navigating through hypertext and it can lead to frustration, loss of interest and decline in efficiency. Therefore, it is plausible to measure the users' perception of the quality of information architecture through perceived disorientation (Ahuja & Webster, 2001). Furthermore, in two psychometric

studies, van Schaik and Ling (2003, 2007) confirmed the reliability, validity and sensitivity of the perceived-disorientation scale, in combination with other interaction experience measures.

The *distraction* and *confusion* sub-categories of the *diversion* category were related to occasions when participants encountered a problem regarding the content, presentation style or information organisation of the site. Therefore, *distraction* and *confusion* need not to be addressed separately in further measurement. However, the *loading time* sub-category represents an additional property of service quality that needs to be accounted for. It is believed that a fast loading time of pages is essential for Web usability (e.g., Nielsen, 1997). A lack of speed in the responsiveness of the interface can seriously undermine interaction experience. Indeed, research shows that the response time of Web sites is an important factor in shaping users' preferences for Web sites (e.g., Chuan-Chuan Lin & Lu, 2000).

The categories that were identified to a large extent represented pragmatic (usability-related) aspects of interaction experience, rather than hedonic (pleasure-related) aspects. This may be due to the nature of interactions that were studied, which can be characterised as 'everyday' experiences rather than (positive) 'peak' experiences. The analysis of interaction-experience questionnaires revealed large and significant differences between regular users of Gazette and non-users. These differences indicate the need to account for the amount of experience and frequency of the participants' use of a particular artefact. In response to this need, future research could have each participant select a news site of his/her own choice that he/she regularly uses and then collect responses to the selected measures. Another advantage of using multiple artefacts in future research is that this way the results

could be generalised to the population of news sites, depending on the selection of news sites that is represented in the sample data, thereby avoiding the fallacy of the product as a fixed effect (Monk, 2004).

Correlation and regression analyses of the interaction-experience questionnaires revealed strong relationships between the scales. *Pragmatic quality* and *hedonic quality* as dimensions of perceived product character were regressed onto the two dimensions of *perceived aesthetics* that represented perceived product characteristics. *Classical aesthetics* was a significant predictor of *pragmatic quality*, and both *classical aesthetics* and *expressive aesthetics* were significant predictors of *hedonic quality*. In turn, *pragmatic quality* was a significant predictor of *goodness*, and *hedonic quality* was a significant predictor of both *beauty* and *goodness* evaluations. The regression analyses support the user-experience model for the application domain of news sites, with the limitation that only perceived aesthetics was used to measure perceived product characteristics. Finally, overall evaluative judgements of the artefact were strongly correlated with the following measures of interaction quality: *intensity of flow* and, particularly, *perceived enjoyment*. A conceptual analysis was conducted to assess the correspondence between the categories and sub-categories of interaction experience that were identified in the analysis of the think-aloud protocols and the psychometric measures that were used in the post-use questionnaire. The results presented in Table 3.17 show that the categories and sub-categories can be related to specific psychometric measures.

Table 3.17. Correspondence between (sub)categories and interaction experience measures.

(Sub)category	Psychometric measure
Impression	Beauty, Goodness, Perceived expressive aesthetics
Content	
Information content	Perceived enjoyment, hedonic quality
Outdated content	Hedonic quality
Advertisement	Intensity of flow
Layout	Pragmatic quality, perceived classical aesthetics, intensity of flow
Information architecture	
Links	Pragmatic quality
Navigation	Pragmatic quality
Structure	Pragmatic quality
Diversion	
Distraction	Intensity of flow
Confusion	Pragmatic quality
Loading time	Intensity of flow

The main objective of the current study was to identify self-reported aspects of experience with a news site. Because the participants of the think-aloud study were not given set behavioural tasks and were not prompted as to which aspects of their experience should they give verbal accounts of, the identified categories can be considered spontaneous, self-reported aspects of experience, which shed light on what may constitute as important factors of interaction experience with news sites. Together with the review of psychological models of technology acceptance and interaction experience presented in the previous chapter, the results of this study were used to select variables from existing research literature to form the basis of a measurement model of interaction experience with news sites. The selection of psychometric measures, data collection for the measurement model and the measurement model itself are presented in the next chapter.

Chapter 4

A measurement model of interaction experience with news sites

Contents

4.1. Overview

This chapter presents a measurement model of interaction experience with news sites. The measurement model (Pedhazur & Pedhazur-Schmelkin, 1991) specifies the relationship between the variables of interaction experience and their observed indicators. Based on the review of models of interaction experience presented in Chapter 2 and the findings of the exploratory study (Study 1) presented in Chapter 3, the current chapter starts with the identification of variables of interaction experience to be included in the measurement model. Existing measurement methods are identified for the variables, drawn from the models discussed in Chapter 2 and from the relevant psychological literature. Next, an online study (Study 2) is presented that was conducted to measure the variables selected for the measurement model. Factor analyses and partial-least-squares structural equation modelling was used to assess the psychometric properties of the variables. The internal consistency, convergent validity, discriminant validity and factor structure of the variables are assessed and presented.

4.2. Selection of measures

4.2.1. Usability

In Chapter 2, usability was defined for the purpose of this thesis and it was identified as an important element of interaction-experience models and research. Research related to each model of interaction experience, presented in Chapter 2, included the conceptualisation and measurement of usability.

The user-experience model describes pragmatic product attributes as properties that are related to the manipulation of interactive artefacts (Hassenzahl, 2003). Pragmatic attributes, such as ‘useful’ and ‘controllable’, allow for accessing the

functionalities of interactive products “to fulfil externally given or internally generated behavioural goals” (Hassenzahl, 2003, p. 34). The pragmatic quality of an interactive product corresponds to its user-perceived usability. Perception of pragmatic attributes was found to be a significant predictor of goodness evaluation of Web sites (van Schaik & Ling, 2008). Perception of the pragmatic quality of interactive products is often measured with 7-point semantic differentials (e.g., Hassenzahl, 2004).

In the environmental-psychology model, usability, as captured subjectively by self-reports, is viewed as an equivalent of *perceived ease of use* in the technology acceptance model (Porat & Tractinsky, in press). In the study that formed the basis of the environmental-psychology model, usability was positively related to the pleasure and dominance dimensions of emotional responses, suggesting that increased ease of system manipulation made interactions more pleasurable and made people feel like more in control. Usability was measured in the study with four items judged on a 7-point Likert scale. The scale conceptualised usability specifically for the use of Web sites with focus on the ease of navigation and ease of use of a particular site. The scale proved to be highly reliable in the referenced study (Cronbach's alpha = .90) and it was also used in previous studies by the authors (Porat & Tractinsky, 2006; Porat et al., 2007).

In the model of user-interface quality assessment, usability is one of the decision-making criteria that form the basis of preference judgements, according to users’ goals and tasks. In the experiments to test the model (Hartmann et al., 2008), usability was assessed in different ways, applying performance analysis, self-reported usability problems with severity rating and a usability scale (Lavie & Tractinsky, 2004). According to the model, if usability is selected as a relevant

criterion for a particular user's goals and task domain in a particular use situation, it may influence interaction experience and the judgement of interaction quality. However, because the usability of news sites is not the focus of the current research and given the broad range of aspects of interaction experience to be included in measurement, it would be time-consuming and inconvenient to ask participants to report on usability problems that they encounter during interaction and rate their severity. Therefore, the use of questionnaires and objective measures such as behavioural metrics (e.g., Juvina & Oostendorp, 2006) is preferred.

The components of user-experience model distinguishes instrumental and non-instrumental product attributes, similar to the distinction between pragmatic and hedonic attributes in the user-experience model. "Instrumental qualities concern the experienced support the system provides and the ease of its use. Features such as the controllability of the system and the effectiveness of its functionality fall into this category" (Thüring & Mahlke, 2007, p. 263). Therefore, the category of instrumental qualities in the components of user-experience model fundamentally corresponds to usefulness and ease of use, that is, usability. In the experiments that formed the basis of the model, Thüring and Mahlke (2007) applied behavioural metrics (e.g., completion rates and time on task) and selected dimensions (controllability, effectiveness, helpfulness and learnability) of the subjective usability measurement inventory (SUMI; Kirakowski & Corbett, 1993) to measure usability. Although it may be informative to assess the role of different dimensions of usability in the formulation of interaction experience, a simpler subjective measurement of usability is preferred for the purpose of the current measurement model to keep the overall number of items at a low level for each measure involved. The attempt to minimise the number

of items per scale is motivated by the large number of aspects of experience to be addressed in the measurement model.

To measure usability conceptualised as user-perceived ease of use, the 4-item pragmatic quality scale of the abridged version of the AttrakDiff2 questionnaire was selected for the purpose of the measurement model (Hassenzahl & Monk, 2010; Hassenzahl et al., 2010). The items of the scale are measured using 7-point semantic differentials. The scale produced satisfactory internal consistency (Cronbach's alpha = .70) in Hassenzahl and colleagues' recent study (2010). Although the usability scale applied by Porat and Tractinsky (in press) had greater internal consistency and its items were specified for the context of Web sites, the pragmatic-quality scale was chosen for the purposes of the measurement model for several reasons. Firstly, the rest of the abridged version of the AttrakDiff2 questionnaire is also proposed to be used for the measurement of hedonic quality and overall product evaluation (see the following sections). Secondly, the psychometric properties of the questionnaire are well-established and conceptually embedded in previous research. Thirdly, the relationships between pragmatic quality, hedonic quality and overall product evaluation (as measured by the abridged AttrakDiff2 questionnaire) with need fulfilment and affective experience has been recently explored by Hassenzahl and colleagues' (2010) research. Therefore, the application of the abridged AttrakDiff2 questionnaire will allow for the comparison of outcomes with existing research.

4.2.2. Non-instrumental characteristics of interaction

As discussed in Chapter 2, models of interaction experience all include non-instrumental aspects of product use. These aspects seem to be important to users,

but they are not connected directly to their performance with a particular system or technology. Non-instrumental interaction characteristics can be useful in predicting future use-behaviour and the evaluation of interactive products.

The perceived capability of interactive systems to provide positive interaction experience can be characterised as hedonic quality. Hedonic quality is "a judgement with regard to a product's potential to support pleasure in use and ownership, that is, the fulfilment of so-called "be-goals" (e.g., to be admired and to be stimulated)" (Hassenzahl et al., 2010, p. 357). In the user-experience model, hedonic (or pleasure-producing) product attributes are constructed by users based on a product's features (perceived product-character) and their personal standards or expectations (Hassenzahl, 2003). Overall product evaluations (beauty and goodness) are influenced by judgements of hedonic quality, which proved to be more stable than pragmatic quality with experience of web-site use (van Schaik & Ling, 2008). In a regression analysis conducted as part of the exploratory study, hedonic quality proved to be the most important predictor of the evaluation of goodness of a news site and a significant predictor of beauty evaluation. Hedonic quality was strongly related to, and could be significantly predicted from, classical and expressive aesthetics evaluation (for details, see Chapter 3).

For the measurement of hedonic quality, the 4-item hedonic quality scale from the abridged version of the AttrakDiff2 questionnaire was adopted (Hassenzahl & Monk, 2010). The items were measured using 7-point semantic differentials. The internal consistency of the scale was satisfactory in recent research (Cronbach's alpha = .75) (Hassenzahl et al., 2010). This scale has been used recently in Hassenzahl and colleagues' study (2010) in relation to the positive and negative affect schedule

(PANAS; Watson et al., 1988) and universal human needs scales (Sheldon et al., 2001), and proved to be strongly related to other variables of interaction experience. The application of this hedonic-quality measure is expected to facilitate the comparison of results of the current study with those of previous studies applying a set of the same measures.

Another aspect of non-instrumental interaction characteristics is the level of enjoyment experienced by users during interaction. Perceived enjoyment is defined as "the extent to which the activity of using computers is perceived to be enjoyable in its own right, apart from any performance consequences that may be anticipated" (Davis et al., 1992, p. 1113). Perceived enjoyment characterises intrinsic motivation (Ryan & Deci, 2000a) in the context of human-technology interaction. It is associated with the use of a particular technology at a particular time; thus it may change dramatically over time and across systems. Sun and Zhang (2008) found that perceived enjoyment has a direct effect on perceived ease of use, which mediates the effect of perceived enjoyment on perceived usefulness. Research suggests that perceived enjoyment also has a direct effect on intention to use (Cyr et al., 2006).

In the analysis of the interaction-experience measures in the exploratory study, *hedonic quality* and *perceived enjoyment* were highly and significantly correlated (r = .84, p < .001) and both were strongly related to *goodness* (r = .89, p < .001 and r = .80, p < .001, respectively). Multiple regression analysis revealed that the effect of *perceived enjoyment* on *goodness* became non-significant (β = .17, t = .95, p > .05) when *hedonic quality* (β = .75, t = 4.31, p < 0.001) was included. Collinearity diagnostics showed no collinearity between the two variables (tolerance = .29, *VIF* =

3.42). These findings suggest that the effect of *perceived enjoyment* on *goodness* was mediated through *hedonic quality*.

Perceived enjoyment is included in the measurement model, because research suggests that it may be an important predictor of interaction outcomes, particularly intention to use. A 3-item perceived enjoyment measure was adopted from Sun and Zhang (2008). The scale had high internal consistency in the referenced research (composite reliability = .98). The items were measured using 7-point semantic differentials.

4.2.3. Affective experience

Users' emotional responses during interaction with technology are an important aspect of human-computer interaction research (Hudlicka, 2003) and thought to be an important determinant of the success of interactive technologies (Norman, 2005). However, emotional responses have multiple components (e.g., subjective feelings and physiological reactions) and they can be studied on various levels of complexity, from automatic, instantaneous affective reactions to reflective emotions as consequences of cognitive appraisal processes (Scherer, 1984). There is a wide range of methods available for the measurement of emotional responses associated with the various components of emotions, for example, self-report rating scales to assess different dimensions of subjective feelings, physiological measures of electrodermal activity to assess bodily reactions and questionnaires to assess dimensions of cognitive appraisals (Mahlke & Minge, 2008).

All models of interaction experience discussed in the literature review address users' emotional responses during the interaction to some extent (see Chapter 2 for details). In the components of user-experience model, emotional responses are

presented as one of the three components of interaction experience, along with the perception of instrumental and non-instrumental qualities. The model proposes the measurement of various components of emotions (subjective feelings, motor expressions and physiological reactions). In the environmental-psychology model, the component of subjective feelings of emotions is measured in the theoretical context of environmental psychology. According to the model, users' emotional responses during the interaction are determined by environmental stimuli, such as the perceived aesthetics and usability of an artefact. Three dimensions of emotional states are measured: pleasure, arousal and control. In turn, emotional responses serve as important determinants of attitudes toward the artefact. In the model of user-interface quality assessment (Hartmann et al., 2008), emotions are not addressed explicitly, although the level of engagement can be considered as a conative aspect of emotional responses. With regards to the user-experience model, the relationship between need fulfilment, the perception of pragmatic and hedonic qualities and affect had been investigated in recent studies. These studies are discussed in more detail below.

Sheldon and colleagues (2001) assessed the salience of 10 candidate universal psychological needs and their relation with event-related affect. Event-related affect was measured on 5-point Likert scales, using the 20-item positive and negative affect schedule (PANAS) (Watson et al., 1988). The PANAS comprises 10 positive and 10 negative adjectives (e.g., excited, alert and nervous). Positive and negative affect scores are computed by averaging the appropriate ratings and the affect-balance score is calculated by subtracting the negative affect score from the positive score. Sheldon and colleagues found that the fulfilment of human needs during satisfying events is strongly associated with high positive and low negative event-related affect.

This relationship was confirmed using both US and South-Korean samples, supporting the universality of basic human needs and their association with affective experience.

Connecting need fulfilment during positive experiences to event-related affect was adopted to the field of human-technology interaction research. Hassenzahl (2008) demonstrated a significant association between need fulfilment and affect, measuring the three basic needs of self-determination theory (Ryan & Deci, 2000b) with Sheldon et al.'s measurement items and using a 10-item abridged version of PANAS on a German sample. In addition, different needs were linked to different qualities of affective experience.

In a more recent study, Hassenzahl and colleagues (2010) found a strong relationship between need fulfilment and positive affect, using a 20-item German version of PANAS rated on a 5-point Likert scale for measuring affective experience. The PANAS scale showed good psychometric properties, with Cronbach's alpha of .83 and .87 for positive and negative affect, respectively, and a small, non-significant inter-scale correlation (r = .07). All needs selected for the study were significantly correlated with positive affect and the average of all needs (general need fulfilment) showed the strongest relation to positive affect, suggesting that need fulfilment in general is more important than the fulfilment of any particular need. Mediation analysis showed that the effect of need fulfilment on hedonic product-quality perceptions was partially mediated through positive affect. In other words, whether a product was deemed hedonic depended on the extent of need fulfilment through positive affect during the interaction. Furthermore, moderated mediation analysis showed that the effect of positive affect on hedonic quality perceptions was only

significant when the affective experience was attributed to the interaction. This finding emphasises the importance of assessing the extent to which users perceive that the interaction with an artefact is accountable for the experienced affect, as opposed to other possible sources of positive affect during the interaction.

In summary, theoretical work on interaction experience suggests the importance of the assessment of affective reactions resulting from the interaction with an artefact. The relationship between need fulfilment and affective experience is well-supported in the literature and has recently been demonstrated in the context of interaction-experience research. The PANAS scale proved to be a useful measurement method in the studies described above. It was applied successfully in different contexts and cultures and it has adequate psychometric properties (Watson et al., 1988). Therefore, the PANAS was selected for the measurement model to measure affective experience.

4.2.4. Need fulfilment

Addressing needs in psychology is plausible, because it may provide an explanation for the motivational basis of a wide variety of behaviours (Sheldon et al., 2001). In the context of interaction experience, the assumption that users' behaviour is driven by underlying needs for certain types of experience justifies the connection between need fulfilment and the quality of experience. Need fulfilment in the course of interaction is expected to yield more positive experiences.

Hassenzahl and colleagues (2010) identified various classifications of needs in the relevant psychological literature and found a relationship between need fulfilment, the perception of hedonic quality and positive affect. The needs of stimulation, relatedness, competence and popularity proved to be especially salient for positive

experiences reported in relation to a wide variety of interactive products. The denotation of pragmatic qualities as hygiene factors and hedonic qualities as motivating factors was also supported. However, the relevant human needs in the case of specific interactive technologies, such as news sites, may be different from the needs that were salient to interactive technologies in general. It is therefore important to establish the relevant list of needs that are addressed by this specific type of application.

Different classifications of needs have been proposed in the psychological literature (see Hassenzahl et al., 2010 for an overview). Sheldon and colleagues (2001) compiled the most comprehensive list of psychological needs derived from a wide range of psychological theories and established their association with positive and negative affect. The list comprises ten needs: autonomy, competence, relatedness, self-actualisation, security, money, influence, physical, self-esteem and pleasure (see Table 4.1 for the description of each need). Fulfilment of the most salient needs showed to be the most strongly associated with high positive and low negative affect. Autonomy, competence and relatedness, all three needs postulated by self-determination theory (Ryan & Deci, 2000b) significantly predicted positive affect in regression analyses. These findings lend support for the assumption that identifying the relevant, salient needs for a particular interactive technology will yield significant predictors of interaction experience.

Based on the exploratory study presented in the previous chapter, the affordances of news sites were identified to derive a tentative set of needs connected to their use. Because stimulation, relatedness, competence and popularity were especially salient in relation to interactive technologies in general (Hassenzahl et al., 2010), these

needs were regarded as tentatively relevant to news sites as well, being a sub-category of interactive technologies. Although *popularity* was among the least salient needs in Sheldon and colleagues' (2001) studies, it would be more relevant to experiences with interactive technologies than positive experiences in general. Although *autonomy* was omitted in Hassenzahl and colleagues' (2010) studies for psychometric considerations, it was retained here, because it was among the most salient needs in Sheldon and colleagues' (2001) studies and it was a significant predictor of positive affect.

Table 4.1. Universal human needs (Sheldon et al., 2001).

Need	Description
Autonomy – independence	Feeling like you are the cause of your own actions rather than feeling that external forces or pressures are the cause of your actions.
Competence – effectance	Feeling that you are very capable and effective in your actions rather than feeling incompetent or ineffective.
Relatedness – belongingness	Feeling that you have regular intimate contact with people who care about you rather than lonely or uncared for.
Self-actualisation – meaning	Feeling that you are developing your best potentials and making life meaningful rather than feeling stagnant and that life does not have much meaning.
Security – control	Feeling safe and in control of your life rather than feeling uncertain and threatened by your circumstances.
Money – luxury	Feeling that you have plenty of money to buy most of what you want rather than feeling like a poor person who has no nice possessions.
Influence – popularity	Feeling that you are liked, respected and have influence over others rather than feeling like a person whose advice or opinions nobody is interested in.
Physical – bodily	Feeling that your body is healthy and well-taken care of rather than feeling out of shape or unhealthy.
Self-esteem – self-respect	Feeling that you are a worthy person who is as good as anyone else rather than feeling like a 'loser'.
Pleasure – stimulation	Feeling that you get plenty of enjoyment and pleasure rather than feeling bored and understimulated by life.

It is proposed that news sites support the fulfilment of the need for *competence-effectance* by allowing for the exploration of online news content. In his influential

paper on human motivation, White (1959) defined competence as “an organism’s capacity to interact effectively with the environment (p. 297)” and described effectance as a tendency to explore and influence the environment. In the context of news sites, the exploration of the environment translates into the exploration of site content (e.g., news, advertisements, blogs and classifieds). It is proposed that feeling competent and effective in one’s actions on a news site is related to the availability of relevant information on the site. For example, if one is interested in the football results, but the football section of a news site fails to provide the relevant summary charts or easy access to them, the browsing episode may result in an unfulfilled information need and a diminished feeling of competence.

An analysis of the screen-capture recordings from the exploratory study revealed that the participants searched for specific topics and pieces of information and explored the site for potentially interesting content. This ‘thematically-driven’ exploration behaviour was more prominent in the group of Gazette readers who, along with the exploration of the site, looked for specific, mainly local-oriented pieces of information (e.g., the state of reconstruction work on a local motorway, market opening times and match results of the local football team). An example of a quote from the think-aloud study (Gazette reader): “Generally, what I’m looking for is local news, local developments ... particularly things like what’s happening on the A66 ... improvement, what’s happening with the local authority” (Participant 3). Non-Gazette readers were predominantly engaged in the exploration of the site and finding potentially interesting content; however, they made specific searches as well (e.g., cricket and football news).

The fulfilment of the need for *stimulation* in the context of news-site use can be characterised as the exploratory behaviour of users on a particular site. During the think-aloud sessions in the exploratory study, all participants were actively involved in browsing for news articles and multimedia information (pictures and videos) of potential interest. An example of a quote from the think-aloud study illustrates this behaviour (Gazette reader): "...and, this is what I normally do, going into local news and seeing what's happened today" (Participant 3). Exploratory behaviour was driven by the titles of articles, article-summaries, content-category hyperlinks (e.g., News, Boro FC, Sports and Blogs) and other hyperlinks (e.g., classifieds, advertisements and picture links) presented on the site. Therefore, it is proposed that the need for stimulation is related to the presentation style or layout and the information architecture of the site. Clear layout and well-presented article summaries are expected to help in focusing users' attention to areas of potential interest. Good information architecture is expected to help users to pursue their interests once they have identified something interesting to read about. Good information architecture should also promote the fulfilment of the need for stimulation by providing an organisation of information under category labels of good information scent. Information scent refers to "the extent to which users can predict what they will find if they pursue a certain path through a Web site" (Nielsen, 2004).

For example, the following quote from a participant in the exploratory study demonstrates how the layout of the page may influence the process of searching for potentially interesting information (non-Gazette reader): "I find the website a bit heavy as well. It's too much information going on and that makes me a bit irritated" (Participant 24). The following quote demonstrates the connection between exploratory behaviour and information architecture (non-Gazette reader): "...and, for

example, if I want to find out a specific topic, for example, like elections or what Gordon Brown said this morning I would know where exactly to look at" (Participant 5). Another quote (non-Gazette reader): "...and then it just brings up a list of the main stories with the headline written in bold, which makes it quite easy to read, so you know which ones you want to look at and which ones you don't" (Participant 11). There was an abundance of references in the units of thought left by the think-aloud participants to being stimulated or bored by certain pieces of content. An example (Gazette-reader): "This is quite interesting ... local interest, history, I find it quite fascinating" (Participant 7). Another example from an under-stimulated user (non-Gazette reader): "Still have a look to see if there's anything interesting. I'm actually thinking to myself this isn't a very interesting website" (Participant 14). Stimulation proved to be a prominent motivation for news-site use in the verbal accounts of think-aloud participants.

It is proposed that news sites support the fulfilment of the needs of *popularity* and *relatedness* through client-to-client facilities, for example, discussion forums and the ability to leave comments and read other users' comments on certain news entries. These facilities allow both for the presentation of one's opinion and the influencing of other people's opinion. Comment functionalities were not used by any of the think-aloud study participants in Study 1, but the results of the online questionnaire indicate their relevance and frequency of use. Although no participant left comments during the think-aloud sessions, 40% (n = 31) of those who completed the online questionnaire reported using the comment functionality of news sites to some extent. Of these participants, 15 (20%) reported commenting *incidentally*, 15 commenting *sometimes* and only 1 participant reported commenting *very frequently*. These findings suggest that the fulfilment of the popularity need might be relevant to a large

proportion of news-site users. Popularity also proved to be relevant to interactive technologies in general in the study of Hassenzahl et al. (2010). Relatedness may also be related to a sense of contact with the community or the feeling of belonging to a community by reading news about it, which may be particularly relevant for the use of local news media (such as the Gazette) and for people's experience of local identity.

In summary, the needs of competence-effectance, pleasure-stimulation, influence-popularity and relatedness-belongingness were identified as relevant for the application domain of online news. These needs were identified based on a review of the literature and an analysis of the tentative affordances of news sites, using the findings of the exploratory study. From the results of the think-aloud study, it appears that the need of pleasure-stimulation is particularly salient for the use of news sites. Although the need for autonomy-independence was not identified during the analysis of news site affordances, it was retained, because it was among the most salient needs in the studies of Sheldon and colleagues (2001).

4.2.5. Interaction outcomes

As discussed in Chapter 2, all models of interaction experience deal with the outcomes of human-computer interaction in various forms, for example, as product evaluations, attitudes toward a product and indicators of experience quality. This (usually) final stage of the models can be regarded as a technology-acceptance component. It is a basic assumption that interaction experience affects technology acceptance. Indeed, the notion that high-quality interaction experience has a positive effect on the acceptance of technologies provides a significant utility to interaction-experience research. If experiences had no effects on users' preferences and

subsequent use behaviour, the usefulness of interaction experience would be very limited. Therefore, it is important to address the measurement of interaction outcomes in a model of interaction experience.

The literature review in Chapter 2 revealed that the outcome variables of technology acceptance and interaction experience models are indicators of interaction experience quality (e.g., pleasure and satisfaction), attitudes toward an artefact, evaluation of an artefact (e.g., beauty and goodness), preference of an artefact over alternative(s), behavioural intention to use and the actual subsequent use of an artefact. Because interaction experience was defined for the purpose of this thesis as an evaluative reflection on one's cognitive-emotional state while interacting with an interactive system, experience quality is regarded as part of the process of interaction, rather than an interaction outcome. Attitudes are often inconsistent with behaviour; they have been removed from the technology acceptance model in order to explain behavioural intention parsimoniously and attitudes were theorised not to be direct determinants of behavioural intention in the unified theory of acceptance and use of technology (Venkatesh et al., 2003). For these reasons, attitudes were omitted from the measurement model. Overall artefact evaluations, however, seem to be very useful outcome variables. For example, beauty and goodness can be measured very conveniently and allow for connecting artefact and interaction characteristics to artefact evaluations. Moreover, overall evaluations of goodness and beauty had been used frequently in research related to the user-experience model, including the exploratory analysis of post-use questionnaires in Study 1. Therefore, the two overall evaluation items of the AttrakDiff2 questionnaire (Hassenzahl, 2004) were adopted for the measurement model.

Behavioural intention can be defined as “the cognitive representation of a decision to perform a given behaviour” (Ajzen, 2002, p. 109). In human-computer interaction, behavioural intention is used in the many versions and applications of the technology acceptance model (e.g., Davis, 1986, Chen & Corkindale, 2008 and Cho et al., 2009) and in the unified theory of acceptance and use of technology (Venkatesh et al., 2003). Intention is assumed to be the immediate antecedent of behaviour and it has been found to be significantly correlated to use (Davis, 1989). Furthermore, because the measurement of the subsequent use of interactive technologies can be extremely hard, it is convenient to predict use behaviour from behavioural intention. Behavioural intention may be regarded as the most direct proxy to actual behaviour. Many authors have developed and used instruments for measuring behavioural intention (e.g., Venkatesh & Davis, 2000). Therefore, items were derived from the existing literature and modified to suit the context of news sites.

To measure the rate of actual use-behaviour, Davis (1989) asked participants to report the degree of current use of the interactive technologies involved in his studies using a 6-point categorical scale (with labels ranging from “Don’t use at all” to “Use several times each day”). He found that the perceived usefulness and perceived ease of use constructs of the technology acceptance model were highly and significantly correlated with use behaviour and that the effect of perceived ease of use was mediated through perceived usefulness. Based on these studies, a 6-point categorical scale was included in the measurement model to assess the self-reported degree of artefact use. (The artefact was a particular news site of the participants’ choice in Study 2.) However, the findings of the online questionnaire in Study 1 suggest that the degree of use of news sites may differ from person to person. Therefore, the degree of use of a particular news site needs to be assessed in

contrast to the individuals' baseline level of news-site use. In order to assess the participants' rate of news-site use, a 6-point categorical scale was adopted from the online questionnaire of Study 1 (see Appendix 3.1, Question 14). Based on these two measures, the rate of use of a particular news site by a particular participant can be assessed relative to his/her rate of news-site use. Consequently, the rate of actual use-behaviour can be compared with the other outcome measures of the measurement model (behavioural intention to use and overall product evaluation) and its relation can be assessed with interaction experience measures.

In summary, measures of behavioural intention to use, overall product evaluations and actual use-behaviour were adopted for the measurement model. Attitude and interaction-quality measures were omitted as outcomes for theoretical considerations. Although subsequent use behaviour is a very informative measure, it is not used in the measurement model, because it can be very hard and laborious to acquire reliable data on it.

4.2.6. Artefact characteristics

Along with need fulfilment, interaction characteristics and interaction outcomes, the categories of units of thought derived from the exploratory study were also considered as artefact factors of interaction experience with news sites. These categories were content, layout, information architecture and diversion[14]. The units of thought that served as the basis for these categories were collected in a setting that aimed to minimise influences from the experimenter and did not include set

[14] The category *impression* is not presented here separately, because it corresponds to overall evaluative judgements already covered by beauty and goodness evaluations that were added to the measurement model as variables of interaction outcomes.

tasks. The researcher did not attempt to elicit any comments about any element of the interaction. It is therefore reasonable to assume that these categories summarise information about the use of a particular news site that were relevant for the participants' experience.

4.2.6.1. Content

The most prominent category in the exploratory study was *content*, with units of thought pertaining to the quality of information presented on the site. Because a news site's most apparent functionality is providing information to its readers, it is important to address the role of information quality in the course of interaction with news sites.

Yang and colleagues (2005) developed and validated an instrument to measure user-perceived service quality of information-presenting Web portals. Information-presenting Web portals are sites that provide online information and information-related services, in contrast to transaction-based or retail-oriented Web sites that focus on online transactions. The interpretation of service quality may differ significantly in the case of these two broad categories of Web sites. For example, the secure and confidential management of credit-card details and the fulfilment of deliveries are of central importance in the case of transaction-based sites, whereas it might not be relevant at all in the case of information-presenting portals such as news sites.

In the discussion of service quality, Yang and colleagues (2005) draw a distinction between information quality and system quality. Information quality is defined as "the users' perception of the quality of information presented on a Web site" (p. 578). Information quality is divided into two dimensions: usefulness of content and

adequacy of information. Usefulness of content refers to the value, reliability, currency and accuracy of information. Specifically, value refers to the relevance and clarity of information and reliability refers to its accuracy, dependability and consistency. Adequacy of information is related to the comprehensiveness of information provided by the site. As opposed to information quality, system quality refers to the users' "perception of a Web site's performance in information retrieval and delivery" (p. 579). This distinction allows for the separation of content from the properties of the system (e.g., navigation and interactivity) in the measurement of service quality.

Information quality can be related to trustworthiness or credibility of news media. According to Kohring and Matthes (2007), trust in news media can be considered a second-order hierarchical factor which consists of four lower order factors, namely: trust in selectivity of topics, trust in selectivity of facts, trust in accuracy of depictions and trust in journalistic assessment. Journalists cannot provide information on every possible issue or event; a selection of covered material is always necessary. The selection of presented content can be based on many different considerations, for example, political views, thematic considerations or presumed relevance for target readers. Trust in the selectivity of topics represents the extent to which readers trust that the reported topics and events are relevant to them and it is addressed in the usefulness-of-content dimension in Yang and colleagues' (2005) measurement dimensions. Trust on selectivity of facts concerns the selection of facts pertaining to the reported topics or events, whereas trust in accuracy of depictions concerns whether the depicted facts are verifiable and truthful. These factors correspond to the accuracy and reliability of information. Trust in journalistic assessment concerns the journalists' commentary on reported topics and events, such as assessment,

opinions and conclusions. Although the issue of journalistic assessment is not addressed in Yang and colleagues' work, it also deals with the comprehensiveness and currency of information. The sub-category *outdated content*, within the category *content*, in the exploratory study supports the need for assessing the currency of information in the assessment of information quality. Kohring and Matthes (2007) developed and validated a measure of trust in news media, but their scale concerns the news media coverage of particular topics and events; therefore, it is inadequate for measuring the trustworthiness of a particular news medium or news site. Online-trust measures are also inadequate, because research in that area typically investigates trust in the context of online retail and e-commerce (Gefen, 2002; Wang & Emurian, 2005). It is therefore plausible to address the conceptualisation of the content category with Yang et al.'s (2005) information quality measure, which conceptually seems to overlap with the dimensions of trust in news media.

Yang et al.'s (2005) usefulness-of-content and adequacy-of-information subscales were adopted for the measurement of the category of content, identified in the exploratory study. The original, 5-point Likert scale measure was developed using the Web site of a commercial property developer; therefore the wording of the items did not fit the application domain of news sites adequately. The relevant items were selected and rephrased for the use of the current study.

4.2.6.2. Layout

The category *layout* in the exploratory study included units of thought on the layout of text and graphic elements on the site and the presentation style of articles. Units of thought in this category typically concerned issues such as the division of text into paragraphs and pages, the placement of certain graphic elements on the pages (e.g.,

advertisements and navigation controls) and the presentation style of individual pages.

User-interface design in Chapter 2 was defined as the structural design of an interface that presents the features and instructional support of an information system. The user-interface acts as a point of contact between users and interactive systems. It reflects the functionalities of a particular system and, in the case of news sites, it allows for accessing information content. Considering the abundance of units of thought regarding page layout in the think-aloud study and the implications of the importance of user-interface design in the case of interactive technologies in human-computer literature (see Chapter 2), it is important to measure perceived user-interface design and establish its connection with other variables of interaction experience.

In an adaptation of the technology acceptance model to Web technologies, Cho and colleagues (2009) developed and established the reliability and the validity of a 4-item measure of perceived user-interface design, which showed to be an important antecedent of continued usage intention in the authors' study. The impact of perceived user-interface design was mediated by perceived functionality and perceived system support, which influenced perceived usefulness and perceived ease of use, respectively. These findings suggest that users make inferences on the usefulness of a particular Web site from the functionality presented on the individual pages of the site. The findings also indicate a connection between perceived user-interface design and perceived ease of use. The 5-point Likert scale measure of perceived user-interface design applied by Cho et al. (2009) was adopted for the purpose of the measurement model.

With regards to layout, the visual appearance or aesthetic quality of the site in question also needs to be considered. Visual aesthetics play an important role in interaction experience studies (e.g., Hassenzahl, 2004; Zhou & Fu, 2007; van Schaik & Ling, 2009) and research suggests that the visual aesthetics of computer interfaces is a strong determinant of users' pleasure and satisfaction (Lavie & Tractinsky, 2004). Visual aesthetics is addressed in some way in all of the interaction experience models discussed in Chapter 2. According to Norman (2005), aesthetics may be even more important in certain cases than usability in users' experience by eliciting a more positive emotional impact of a particular design. Furthermore, research suggests that enhanced aesthetics design can improve task outcome under the condition of poor usability (Moshagen et al. 2009). In the exploratory study, the classical and expressive dimensions of the perceived aesthetics scale (Lavie & Tractinsky, 2004) showed high and significant correlations with measures of perceived enjoyment, intensity of flow, pragmatic quality, hedonic quality and overall beauty and goodness. Therefore, it is reasonable to include a perceived aesthetics measure in the measurement model.

The two-dimension aesthetics questionnaire elicits high-level perceptions of design quality, leading to overall aesthetic evaluations on the two separate aesthetics dimensions (Hartmann et al., 2008). To complement this approach of measuring Web site aesthetics, Hartmann and colleagues (2008) also asked their participants to evaluate the quality of individual design features, applying heuristics of attractiveness (Sutcliffe, 2002). However, they did not report the analysis of the heuristics of attractiveness measures. The two-dimension aesthetics questionnaire, on the other hand, differentiated well between distinct Web site designs and expressive aesthetics proved to be a significant predictor of overall preference. An 8-item version of this

measure (four items for each dimension) with adequate psychometric properties was introduced by Porat and Tractinsky (in press). The 8-item perceived aesthetics scale, using a 7-point Likert scale, was selected for the purpose of the measurement model.

4.2.6.3. Information architecture

The category *information architecture* in the exploratory study comprised units of thought regarding links, navigation and the structure of the site. The *links* sub-category contained units of thought regarding hyperlinks and the connectivity of the site with other Web sites (e.g., Facebook and Twitter). The *navigation* sub-category contained units of thought regarding the navigation between the pages of the site and comments on the search functionality. The *structure* sub-category contained unit of thought regarding the organisation of information on the site.

Large sites containing hundreds of individual pages are complex, richly inter-connected and cross-referenced bodies of multimedia information (Utting & Yankelovich, 1989). News sites may contain several hundred individual pages. Therefore, without proper organisation of content and navigational aids, finding information can be extremely difficult. News entries and other information are organised under distinct sections by various criteria (e.g., novelty, content and location). News sites are updated frequently to provide up-to-date information to their readers. As a result, a specific piece of information can change its location on a site over time. These characteristics of news sites emphasise the importance of clear and intelligible content organisation, and the need to provide search facilities. According to McDonald and Stevenson (1998), disorientation, the tendency to lose the sense of location on a Web site, is one of the most common problems faced by

users navigating through hypertext and it can lead to frustration, loss of interest and decline in efficiency. Disorientation is more accurately defined as "the situation in which the user does not have a clear conception of relationships within the system, does not know his present location in the system relative to the display structure and finds it difficult to decide where to look next within the system" (Woods, 1984, as cited in Ahuja & Webster, 2001, p. 20). The *information architecture* category summarised the issues of navigating the complex structure of hypertext on a news site. Therefore, it is plausible to approach users' perception of the quality of information architecture through disorientation.

For the measurement model, the perceived disorientation scale (Ahuja & Webster, 2001) was selected to address the information architecture category of the exploratory study. Ahuja and Webster (2001) provided experimental support for the idea that disorientation is different from, although strongly and negatively correlated to, the ease of use construct. In their experiment, perceived disorientation was related negatively to information search performance and showed to predict performance better than users' actions (that is, the number of pages opened and the number of pages opened more than once). The perceived measure may have an advantage over the behavioural measure, because navigational actions might be influenced by factors other than disorientation, such as exploratory behaviour (Smith, 1996). In the exploratory study, where no set tasks were given to the participants, exploratory behaviour was prominent in the participants' use of the news site. For this reason, the perceived disorientation measure seems especially plausible for the measurement of disorientation to assess users' perceptions of information-architecture quality of news sites.

Perceived disorientation is measured using seven items rated on a 7-point Likert scale, ranging from ‘never’ to ‘always’. The items include statements such as ‘I felt lost’ and ‘Navigating between the pages was a problem’. The scale showed high reliability (Cronbach's alpha of .90), as well as satisfactory content validity, discriminant- and concurrent validity, and predictive efficacy (Ahuja & Webster, 2001). In two psychometric studies, van Schaik and Ling (2003, 2007) successfully confirmed the reliability, validity and sensitivity of the perceived disorientation scale, in combination with other interaction experience measures, using both a Likert scale and a visual analogue scale as response formats.

4.2.4.3. Diversion

The category *diversion* in the exploratory study was derived from units of thought that suggested the participants’ diversion from their intended interaction with the site. The category was divided into three sub-categories, namely distraction, confusion and loading time.

A unit of thought was assigned to the *distraction* sub-category whenever a participant suggested being distracted by something. These units of thought predominantly regarded distracting graphic elements (layout) and distracting content, mainly advertisements (content). An example of a quote (non-Gazette reader): “...and those moving adverts ... really distracting. I don’t like them at all [...]” (Participant 15). The *distraction* sub-category can be regarded as consisting of quotes that represent layout and content issues. Therefore, there is no need to introduce a separate measure for distraction in the measurement model.

Similar to distraction, the *confusion* sub-category seems to be consisting of units of thought that represent issues with another category, namely, layout. Units of thought

in the *confusion* sub-category generally suggest confusion in relation to the presentation style of the site and visual impressions. An example of a quote (non-Gazette reader): “yeah, it’s very colourful, therefore ... which makes it very confusing to me” (Participant 5). Another example (non-Gazette reader): “My first reaction’s there’s an awful lot of information and I’m really not sure where I’d go first” (Participant 12). These units of thought represent issues with the layout and aesthetics of the page. Measures for these issues were introduced previously in relation to the *layout* category. Another tentatively plausible source of confusion is navigation issues. Although no units of thought were collected during the exploratory study about participants feeling ‘lost’ on the site, disorientation can be a serious issue to influence interaction experience. Disorientation was addressed with the measure *perceived disorientation* in relation to the information architecture category. Therefore, *confusion* can be covered by measures selected to address the *layout* and *information architecture* categories.

The third subcategory of the *diversion* category was *loading time*, comprising both positive and negative units of thought about the time required to load individual pages. Loading time or response time of a particular Web site refers to the time users spend waiting to interact with the site (Chuan-Chuan Lin & Lu, 2000). Studies show that fast response times are essential for Web usability (e.g., Nielsen, 1997). A lack of speed in the responsiveness of the interface can seriously undermine interaction experience. Research shows that the response time of Web sites is an important factor in shaping users’ Web-site preferences (e.g., Chuan-Chuan Lin & Lu, 2000). *Loading time* can be characterised as part of the accessibility of content, which involves two aspects: availability and responsiveness (Yang et al., 2005). It is expected that the content of a particular news site should be available at all times.

Unavailable content and broken links are expected to negatively influence interaction experience and the evaluation of the news site. Yang and colleagues' (2005) Web-site quality measure included a short, two-item scale for accessibility, using a 5-point Likert scale. This accessibility measure was adopted to measure accessibility in Study 2. Both loading time and availability can be considered as a part of a Web site's service quality. Although other aspects of service quality are present in the literature (e.g., security and system support), the literature review and the exploratory study did not indicate a need to include more service-quality variables for news sites.

4.3. Method of Study 2

4.3.1. Design

An on-line interaction experience questionnaire was designed to collect responses to items as indicators of variables measuring aspects of users' interaction experience with news sites. The questionnaire was advertised though university newsletters and student e-mail lists at Teesside University. A prize-draw of £50 was used as an incentive, but psychology students at Teesside University also received research-participation credits for the completion of the questionnaire. Several other British universities were contacted to help in distributing an advertisement containing the link to the study among their students and staff. Answers were collected from Bangor University, City University London, Kingston University and Teesside University.

4.3.2. Materials and procedure

Survey Monkey (*http://www.surveymonkey.com*), an online survey tool was used to design the questionnaire and collect data[15]. Informed consent was collected after the

[15] The questionnaire was piloted with seven postgraduate researchers at Teesside University as respondents.

introduction by requiring participants to tick a series of checkboxes in order to carry on to a set of instructions. Following the consent form, participants were instructed to use a news site of their own choice before proceeding to the questions. At the end of the questionnaire, participants were asked to provide their e-mail address to enter the prize-draw. The complete questionnaire, including the measures discussed in Section 4.2 (with item-codes), is presented in Appendix 4.1. A summary of psychometric measures adopted for the study is presented in Table 4.2.

4.3.3. Participants

Participants had to be over 18 and fluent in English to be eligible for the study. Three hundred and five gave a full response to the interaction-experience questionnaire (120 male, 185 female; mean age = 24.63 years, *SD* = 7.52). The average experience of Internet use was 9.66 years (*SD* = 3.19). Nearly two thirds of the participants used the BBC news site before completing the questionnaire. The majority (94%) accessed the Internet on a daily or more frequent basis and 70% reported to use the Internet for an hour or more per day. Nearly half (47%) reported daily or more frequent access to news sites and 56% used these sites for between 10 and 15 minutes duration per visit. Participants mainly used laptop computers (86%), desktop computers (47%) and mobile phones (31%) to access news sites. Demographics, Internet-use behaviour and news-site use behaviour information are presented in Table 4.3.

Table 4.2. A summary of psychometric measures adopted for Study 2.

Source	Instrument	Dimensions	Number of items
Watson et al., 1988	PANAS	Positive affect	10
		Negative affect	10
Hassenzahl and Monk, 2010	AttrakDiff2	Pragmatic quality	4
		Hedonic quality	4
		Beauty	1
		Goodness	1
Sun and Zhang, 2008	Perceived enjoyment	N/A	3
Ahuja and Webster, 2001	Perceived disorientation	N/A	7
Cho et al., 2009	Perceived user-interface design	N/A	4
Porat and Tractinsky, in press	Perceived aesthetics	Classical aesthetics	4
		Expressive aesthetics	4
Yang et al., 2005	Information quality	Usefulness of content	3
		Adequacy of information	3
	Service quality	Accessibility	2
Venkatesh and Davis, 2000	Behavioural intention	N/A	2
Sheldon et al., 2001	Need fulfilment	Autonomy	3
		Competence	3
		Relatedness	3
		Stimulation	3
		Popularity	3

Note. All scales were measured using 7-point Likert scales, except for AttrakDiff2, which was measured with 7-point semantic differentials.

Table 4.3. Descriptive statistics of demographics, Internet-use behaviour and news-site use behaviour.

Age	Mean	SD	Minimum	Maximum	N
	24.63	7.52	18	62	305
Gender		Male		Female	
		120 (39%)		185 (61%)	
News site used (during study)	BBC	The Guardian	Sky News	The Independent	Other
	202 (66%)	18 (6%)	12 (4%)	5 (2%)	68 (22%)
Internet use (years)			Mean	*SD*	
			9.66	3.19	
Internet use frequency			*N*		Percentage
	Weekly		3		1.0
	2-3 times a week		5		1.6
	4-6 times a week		9		3.0
	Daily		65		21.3
	2-3 times a day		51		16.7
	More than 3 times a day		172		56.4
Internet use duration per visit					
	1-10 minutes		5		1.6
	About 15 minutes		15		4.9
	About 30 minutes		43		14.1
	About 45 minutes		29		9.5
	About 1 hour		73		23.9
	Several hours		140		45.9
News-site use frequency					
	Weekly		86		28.2
	2-3 times a week		43		14.1
	4-6 times a week		34		11.1
	Daily		103		33.8
	2-3 times a day		24		7.9
	More than 3 times a day		15		4.9
News-site use duration per visit					
	1-4 minutes		26		8.5
	About 5 minutes		43		14.1
	About 10 minutes		74		24.3
	About 15 minutes		96		31.5
	About 30 minutes		48		15.7
	More than 30 minutes		18		5.9
Devices used to access news sites					
	Desktop computer		142		46.6
	Laptop computer		261		85.6
	Mobile phone		94		30.8
	PDA		4		1.3
	Other[a]		13		4.3

[a]Other devices used to access news sites were: iPad, iPod touch, games console, tablet PC and Kindle.

4.4. Analysis

4.4.1. Analysis method

Along with factor analysis, for the formulation of a measurement model of interaction experience with news sites (and for subsequent structural models in Chapters 5 and 6) partial-least-squares path modelling (PLS) was used for the following reasons (see Vinzi et al., 2010). PLS allows for the integrated analysis of a measurement model, which specifies the relationships between latent variables and their manifest variables, and a structural model, which specifies the relationships between latent variables. PLS has less stringent assumptions regarding the distribution of variables and error terms than covariance-based structural equation modelling, and supports both reflective and formative measurement. The sample-size requirement for PLS is also lower than required for covariance-based structural equation modelling. PLS maximises the explained variance in dependent variables and it is suitable for estimating complex models (multi-stage models with a high number of latent variables and connections); therefore, it is adequate for prediction-oriented research involving a wide range of variables. Latent variable scores in PLS are exact linear combinations of manifest variables, rather than average scores of manifest variables calculated for each latent variable with satisfactory internal consistency. Therefore, PLS provides more accurate scale values than the technique of averaging item scores. All PLS analyses in this thesis were conducted using the SmartPLS software (*http://www.smartpls.de*). During PLS analyses in this thesis, bootstrapping samples of 5000 were used to test the significance of model parameters, as recommended by Henseler et al. (2009).

4.4.2. Factor structure of measures

A series of factor analyses and PLS analyses with factor weighting scheme (Chin, 2010) were conducted in order to examine the factor structure of the measures before specifying a PLS structural model. The perceived product-characteristic measures of *perceived aesthetics*, *perceived disorientation*, *perceived user-interface design*, *usefulness of content*, *adequacy of information* and *accessibility* were analysed together, followed by the analysis of the po*sitive and negative affect schedule*, *AttrakDiff2* (abridged version) and *perceived enjoyment* scales as interaction-experience measures[16]. Following this, the factor structure of the five selected need-fulfilment subscales is presented and discussed. Finally, a general measurement model of all measures that were selected for structural modelling is presented.

4.4.2.1. Factor structure of measures of perceived product characteristics

For perceived aesthetics, principal component extraction was used with the explicit extraction of 2 factors (for classical and expressive aesthetics) and varimax rotation (*KMO* = .90; Bartlett: $\chi^2(28) = 1321.71$, $p < 0.001$; total variance extracted: 70%). Only one principal component had an eigenvalue over 1 and the rotated solution failed to reproduce the original factor structure of the measure. Furthermore, the rotated solution contained several high cross-loadings and simple structure was not achieved. Oblique rotation (direct oblimin) led to a similar structure. However, a single-factor solution with 60% of extracted variance resulted in a simple structure. All items loaded highly on one component and the resulting scale showed

[16] Measures of perceived product characteristics and measures of interaction experience were analysed separately using factor analysis to explore the factor structure of these measures in detail before presenting a overall measurement model including all items of all measures.

satisfactory internal consistency (Cronbach's alpha = .90). The item-loadings of the factor solutions are presented in Table 4.4.

Table 4.4. Component matrices of perceived-aesthetics items.

	Factor (varimax rotation)		Factor (direct oblimin rotation[a])		Single-factor
	1	2	1	2	solution
Sophisticated	.84	.14	.92	-.16	.77
Creative	.76	.39	.79	.14	.85
Clean	.76	.11	.82	-.16	.67
Aesthetic	.72	.44	.74	.20	.84
Pleasant	.67	.44	.67	.23	.80
Original	.63	.54	.62	.34	.83
Symmetrical	.09	.87	.-04	.89	.58
Spectacular	.45	.75	.38	.64	.81

Note. Extraction method: principal components.

[a]The pattern matrix is presented; loadings are regression coefficients.

In order to determine whether a two-factor or a single-factor solution should be used for aesthetics in the measurement model, the factor structure of perceived aesthetics was examined using PLS. SmartPLS can only calculate PLS equations on well-formed structural models; therefore, case number was used as dependent variable, because it was expected (and proved) to be uncorrelated with the two perceived aesthetics subscales. The subscales were assigned their original items (four each). The item loadings and cross-loadings are presented in Table 4.5.

Table 4.5. Loadings and cross-loadings of perceived-aesthetics items.

Latent variable	Item	Factor		Single factor
		1	2	solution
Classical aesthetics	Clean	0.72	0.52	0.64
	Pleasant	0.80	0.70	0.78
AVE = .58	Symmetrical	0.55	0.47	0.53
	Aesthetic	0.93	0.75	0.87
Expressive aesthetics	Original	0.68	0.90	0.84
	Sophisticated	0.62	0.78	0.75
AVE = .67	Spectacular	0.65	0.75	0.74
	Creative	0.74	0.84	0.84

Note. Latent-variable analysis through PLS.

All perceived aesthetics items produced high cross-loadings on both subscales. The latent variables *classical aesthetics* and *expressive aesthetics* were highly correlated (r = .80). The square of the correlation between the two subscales (r^2 = .64) was larger than the average variance extracted (*AVE*) of the classical aesthetics subscale (*AVE* = .58), which indicates a lack of discriminant validity, according to the Fornell-Larcker criterion (Fornell & Larcker, 1981). However, a combined *perceived aesthetics* latent variable with all items included from both subscales retained a satisfactory amount of variance from the items (*AVE* = .57) and exhibited satisfactory internal consistency (composite reliability = .91). Therefore, *perceived aesthetics* was treated as a single factor measure in further analyses, measured by a composition of the four classical and four expressive aesthetics items.

The measures of *usefulness of content* (UC), *adequacy of information* (AI) and *accessibility* (AC) were adopted from the same instrument (Yang et al., 2005), where the factor structure was established using factor analysis with principal component extraction and varimax rotation. Therefore, the same analysis was used in the current study, with the explicit extraction of three factors (*KMO* = .84; Bartlett: $\chi^2(28) = 1040.06$, $p < 0.001$; total variance explained: 75%). The rotated component matrix is presented in Table 4.6.

Table 4.6. Rotated component matrix of all usefulness of content, adequacy of information and accessibility items.

		Factor		
		1	2	3
UC1	The site provides relevant information.	.87	.14	.22
UC2	The site provides up-to-date information.	.83	.18	.21
UC3	The site provides unique content.	-.01	.86	-.05
AI1	The site provides comprehensive information.	.37	.68	.33
AI2	The site provides complete content.	.43	.68	.27
AI3	The site provides sufficient information.	.52	.43	.41
AC1	The pages of the site are accessible.	.41	.19	.77
AC2	The pages of the site load quickly.	.14	.07	.91

Note. Extraction method: principal component. Rotation: varimax. UC: usefulness of content. AI: adequacy of information. AC: accessibility.

Accessibility reproduced well, with both its items loading highly on one factor. There were two problematic items: UC3 loaded highly on the factor containing the items of *adequacy of information* without loading on the factor containing the other two *usefulness-of-content* items, and AI3 loaded on all three factors. These items were removed and the same analysis was repeated with two items for each scale (*KMO* = .81; Bartlett: $\chi^2(15) = 741.51$, $p < 0.001$; total variance explained: 84%). The analysis resulted in a simple structure (see Table 4.7); therefore, this solution was used in further analyses.

Table 4.7. Rotated component matrix of selected usefulness of content, adequacy of information and accessibility items.

		Factor		
		1	2	3
UC1	The site provides relevant information.	.25	.86	.22
UC2	The site provides up-to-date information.	.26	.85	.23
AI1	The site provides comprehensive information.	.85	.24	.23
AI2	The site provides complete content.	.87	.25	.15
AC1	The pages of the site are accessible.	.34	.36	.74
AC2	The pages of the site load quickly.	.12	.16	.93

Note. Extraction method: principal components. Rotation: varimax

In order to analyse the factor structure of the measures of perceived product characteristics together, the items of all measures were entered a single factor analysis with principal component extraction method and varimax rotation. Seven factors were extracted (*KMO* = .91; Bartlett: $\chi^2(325) = 4369.01$, $p < 0.001$; total variance explained: 75%), one for each measure: *classical aesthetics* (CA), *expressive aesthetics* (EA), *perceived disorientation* (PD), perce*ived user-interface design* (PUID), *usefulness of content* (UC), *adequacy of information* (AI) and *accessibility* (AC). The rotated factor matrix is presented in Table 4.8.

Table 4.8. Rotated factor matrix of items of all measures of perceived product characteristics.

	Factors						
	1	2	3	4	5	6	7
PD6	.84						
PD4	.84						
PD2	.79						
PD7	.78						
PD5	.77						
PD3	.74						
PD1	.69		-.32				
EA4		.85					
EA2		.80					
CA4		.80					
EA1		.75				.33	
EA3		.70				.30	.43
CA1		.61		.36			
CA2		.61					
PUID1			.83				
PUID2		.30	.82				
PUID3	-.32	.32	.77				
UC2	-.30			.78			
UC1				.77			
AC2	-.38				.75		
AC1	-.33			.31	.70		
AI3				.47	.59		
AI1				.45	.47	.42	
UC3		.33				.81	
AI2				.46	.46	.46	
CA3		.36					.87

Note. Extraction method: principal components. Rotation: varimax. Loadings < .30 suppressed.

The scales *perceived disorientation* (Factor 1) and *perceived user-interface design* (Factor 3) reproduced well in the factor solution. In accordance with previous analyses, both *perceived aesthetics* subscales loaded on the same factor (Factor 2), which was clearly discernible from the others. The items of *usefulness of content*, *adequacy of information* and *accessibility* showed high cross-loadings. However, this was expected, because the problematic items of UC3 and AI3 were retained in this analysis. The factor *adequacy of information* did not reproduce clearly in this analysis.

4.4.2.2. Factor structure of interaction experience measures

For the assessment of the factor structure of the positive and negative affect schedule (PANAS), principal component extraction method was used, with the explicit extraction of two factors, and varimax rotation (see Table 4.9). The factor solution was a moderate fit (*KMO* = .86; Bartlett: $\chi^2(190) = 2151.59$, $p < 0.001$; total variance explained: 46%), because of the relatively low amount of explained variance. There were four main components with eigenvalue greater than 1, but the scree-plot indicated the plausibility of the extraction of two factors, because of the drop of eigenvalues after the second component (eigenvalues: 1. = 5.44, 2. = 3.85, 3. = 1.19 and 4. = 1.07). Nevertheless, simple structure was achieved with all positive-affect items loading on one factor and all the negative-affect items loading on the other. All cross-loadings were lower than .30 and the scale inter-correlation was low ($r = .20$, $p < .01$), suggesting good discriminant validity. The internal consistencies of the positive and negative scales were satisfactory (Cronbach's alphas .87 and .88, respectively).

Table 4.9. Rotated component matrix of PANAS items.

Scale	Item	Factor 1	Factor 2
Positive affect	Inspired	.81	.01
	Enthusiastic	.75	-.02
	Determined	.75	.18
	Active	.74	.08
	Proud	.71	.08
	Excited	.69	.07
	Attentive	.66	.04
	Strong	.60	.15
	Alert	.56	.16
	Interested	.55	-.17
Negative affect	Distressed	-.06	.76
	Afraid	-.01	.74
	Upset	.03	.70
	Scared	.14	.69
	Ashamed	.05	.68
	Nervous	.08	.67
	Jittery	.18	.61
	Hostile	.18	.60
	Irritable	-.00	.59
	Guilty	.01	.54

Note. Extraction method: principal components. Rotation: varimax.

Hassenzahl et al. (2010) used principal component analysis with varimax rotation and the explicit extraction of two factors to test the factor structure of the abridged AttrakDiff2 questionnaire. Therefore, the same analysis was carried out on the items of the *pragmatic quality* and *hedonic quality* scales (4 items each). The analysis resulted in a satisfactory solution (*KMO* = .81; Bartlett: $\chi^2(28) = 695.22$, $p < 0.001$; total variance explained: 59%), with items loading on the appropriate factors (see Table 4.10).

The internal consistencies of the *pragmatic quality* and *hedonic quality* scales were satisfactory (Cronbach's alphas .70 and .78, respectively). Scale inter-correlation was moderate ($r = .45$, $p < .01$), but smaller than the internal consistencies, which indicates discriminant validity. (Inter-correlation between the scales is presented together with coefficients of discriminant validity in Table 4.14).

Table 4.10. Rotated component matrix of AttrakDiff2 items.

Scale	Item	Factor	
		1	2
Hedonic quality	Unimaginative - Creative	.81	.03
	Tacky - Stylish	.77	.19
	Dull - Captivating	.74	.29
	Cheap - Premium	.71	.10
Pragmatic quality	Impractical - Practical	.28	.77
	Confusing - Structured	.30	.75
	Unpredictable - Predictable	-.17	.67
	Complicated - Simple	.27	.66

Note. Extraction method: principal components. Rotation: varimax.

The interaction-experience measures of *positive affect*, *negative affect*, *pragmatic quality*, *hedonic quality* and *perceived enjoyment* (PE) were analysed in a single analysis with principal component extraction and varimax rotation, with the explicit extraction of five components (one for each scale). The analysis resulted in a satisfactory solution (*KMO* = .87; Bartlett: $\chi^2(465) = 3616.97$, $p < 0.001$; total variance explained: 55%). The loadings of each item are presented in Table 4.11.

All interaction-experience measures essentially reproduced in the combined factor analysis. *Negative affect* reproduced without any high cross-loadings on other factors, whereas the item ‘interested’ from *positive affect* loaded higher on the factor of *perceived enjoyment*. All items of *perceived enjoyment* cross-loaded on the factor of *hedonic quality*, but the loadings were nearly twice as large as the cross-loadings for each item. Finally, the item ‘complicated - simple’ of *pragmatic quality* cross-loaded highly on the factor of *hedonic quality*. In agreement with the separate factor analyses, the scales were kept in their original form for further analysis.

Table 4.11. Rotated component matrix of items of interaction-experience measures.

Scale	Item	Factor				
		1	2	3	4	5
Positive affect	Determined	.80				
	Inspired	.76				
	Proud	.75				
	Enthusiastic	.71				
	Active	.70				
	Strong	.67				
	Excited	.66				
	Attentive	.52			.35	.35
	Alert	.50				
	Interested	.33			.46	
Negative affect	Afraid		.77			
	Scared		.73			
	Distressed		.72			
	Upset		.71			
	Ashamed		.68			
	Nervous		.67			
	Jittery		.59			
	Hostile		.59			
	Irritable		.56			
	Guilty		.56			
Hedonic quality	Tacky - Stylish			.70		
	Cheap - Premium			.69		
	Unimaginative - Creative			.68		
	Dull - Captivating			.67		
Perceived enjoyment	PE1			.30	.78	
	PE2			.37	.75	
	PE3	.37		.35	.68	
Pragmatic quality	Impractical - Practical			.31		.64
	Unpredictable - Predictable					.62
	Confusing - Structured			.34	.32	.57
	Complicated - Simple			.48		.50

Note. Extraction method: principal components. Rotation: varimax. Loadings < .30 suppressed.

4.4.2.3. Factor structure of need-fulfilment subscales

Five of the 10 universal human needs proposed by Sheldon and colleagues (2001) were selected for the study previously in this chapter: autonomy, competence, relatedness, popularity and stimulation. Similar to the original study of Sheldon et al. (2001) and that of Hassenzahl et al. (2010), principal component extraction with varimax rotation was used to identify the underlying factor structure of the selected

items (3 items each). Only two components had eigenvalues greater than 1 and the resulting solution with the explicit extraction of five components failed to reproduce, or even to resemble, the original factor structure (see Table 4.12). Direct oblimin rotation yielded very similar results, and a two-factor solution based on the eigenvalues did not result in an interpretable structure.

Table 4.12. Rotated factor matrix of the selected need-fulfilment scales.

	Factor				
Item	1	2	3	4	5
Popularity 2	.86				
Popularity 3	.84				
Relatedness 1	.83	.32			
Relatedness 3	.82				
Relatedness 2	.81				
Competence 2	.78		.37		
Popularity 1	.72			.32	
Competence 1	.67		.31		.35
Stimulation 1	.62		.57		
Autonomy 1		.80			
Autonomy 3	.47	.68			
Stimulation 3	.38		.73		
Stimulation 2	.48	.40	.56		
Competence 3				.88	
Autonomy 2				.30	.86

Note. Extraction method: principal components. Rotation: varimax. Loadings < .30 suppressed.

Both Sheldon and colleagues and Hassenzahl and colleagues asked their participants to recall ‘peak experiences ’. However, in the current study, need fulfilment was assessed after use and the participants were asked to report on the everyday-experience of browsing a news site. It may be that the fulfilment of universal needs, as measured by these scales, is less salient in less satisfying, but nonetheless pleasurable or otherwise rewarding, experiences. It may also be that the statements of the scales are more suitable for the description of outstanding or extraordinarily satisfying experiences, and they do not apply to more common experiences, such as the use of news sites. Finally, in the aforementioned previous

studies, participants were asked to report on an exceptional experience of their past (cumulative account of experiences), whereas in the current study they commented on a 'fresh' experience (episodic account of an experience). The immediate reflection on a common experience may differ from the recollection of a memory of an outstanding experience. Because of the unsuccessful reproduction of their original factor structure, the need-fulfilment scales were excluded from further analysis.

4.4.3. A measurement model of all measures selected for structural modelling

In line with the suggestion of Chin (2010), a measurement model of all included scales was tested using PLS by drawing all possible structural links between the constructs, with the PLS inner weighting option set to factorial scheme in the SmartPLS analysis software. The obtained coefficients of reliability and convergent validity are presented in Table 4.13, coefficients of discriminant validity are presented in Table 4.14 and item cross-loadings are presented in Appendix 4.2.

Of the items of *pragmatic quality*, PQ2 (unpredictable - predictable) had a low loading (.39). When excluded from the analysis, the average extracted variance of the *pragmatic quality* scale increased to .66 from .53, and the scale's composite reliability (*CR*) increased to .85 from .81. However, the sale's composite reliability and average extracted variance were still acceptable with PQ2 included. Furthermore, the examination of the cross-loadings of PQ2 revealed that it did not load highly on any other scales and its second highest loading after *pragmatic quality* was on *hedonic quality* with 0.15 (see Appendix 4.2). Therefore, PQ2 was retained in order to facilitate comparability with previous studies.

Table 4.13. Coefficients of reliability and convergent validity.

Construct/indicator	Average variance extracted	Composite reliability	Loading	Standard error	t^a
Perceived aesthetics	0.58	0.92			
CA1			*0.64	0.05	12.87
CA2			0.83	0.02	46.27
CA3			*0.58	0.05	10.53
CA4			0.85	0.02	49.36
EA1			0.80	0.02	32.09
EA2			0.72	0.04	16.16
EA3			0.78	0.03	29.20
EA4			0.85	0.02	50.40
Perceived disorientation	0.69	0.94			
PD1			0.88	0.02	51.93
PD2			0.85	0.03	31.10
PD3			0.73	0.04	16.67
PD4			0.81	0.03	23.45
PD5			0.85	0.02	35.20
PD6			0.89	0.02	43.42
PD7			0.80	0.03	24.46
Perceived user-interface design	0.87	0.95			
PUID1			0.92	0.01	70.38
PUID2			0.95	0.01	130.37
PUID3			0.93	0.01	86.83
Usefulness of content	0.85	0.92			
UC1			0.93	0.01	64.68
UC2			0.90	0.03	31.87
Adequacy of information	0.85	0.92			
AI1			0.92	0.01	62.96
AI2			0.92	0.01	66.88
Accessibility	0.84	0.92			
AC1			0.92	0.02	59.96
AC2			0.92	0.02	51.93
Pragmatic quality	0.53	0.81			
PQ1			0.85	0.02	42.02
PQ2			*0.39	0.09	4.57
PQ3			0.81	0.04	21.20
PQ4			0.76	0.03	21.81
Hedonic quality	0.60	0.86			
HQ1			0.80	0.03	30.97
HQ2			0.80	0.03	31.52
HQ3			*0.69	0.06	11.13
HQ4			0.81	0.03	28.35
Perceived enjoyment	0.82	0.93			
PE1			0.92	0.01	77.86
PE2			0.92	0.01	67.50
PE3			0.88	0.01	58.13
Positive affect	0.46	0.89			
POS1			*0.64	0.04	15.25
POS2			*0.68	0.04	15.84

POS3			*0.58	0.05	11.56
POS4			0.71	0.04	19.30
POS5			*0.68	0.04	15.99
POS6			*0.59	0.05	10.85
POS7			0.78	0.02	32.19
POS8			0.71	0.04	17.41
POS9			*0.68	0.04	15.37
POS10			0.70	0.04	19.12
Negative affect	0.46	0.89			
NEG1			*0.66	0.06	10.23
NEG2			*0.56	0.09	6.53
NEG3			*0.67	0.06	10.60
NEG4			*0.67	0.08	8.33
NEG5			*0.66	0.06	10.59
NEG6			*0.69	0.05	14.01
NEG7			0.74	0.05	14.27
NEG8			0.70	0.05	13.19
NEG9			0.72	0.05	13.14
NEG10			0.70	0.07	9.88
Behavioural intention	0.94	0.97			
BI1			0.97	0.01	129.60
BI2			0.97	0.01	79.72

[a]Bootstrap, *N* = 5000

*Loading < 0.70.

AVE values of the scales *positive affect* and *negative affect* did not reach.50 (recommended by Chin, 2010). However, all items loaded significantly on the corresponding scales and loadings were consistently and markedly higher than cross-loadings (see Appendix 4.2.); therefore, construct validity at the item level was supported for both affect dimensions. In the rest of the scales, values of average variance extracted exceeded .50, meaning that more than 50% of variance of indicators was accounted for by the scales. Composite reliability (CR) values were larger than .80 for each scale, indicating satisfactory internal consistency.

Table 4.14. Coefficients of discriminant validity and inter-correlation between the scales.

	AC	AES	AI	ATT	BUSE	BEAU	BI	GOOD	HQ	NEG	PD	PE	POS	PQ	PUID	UC	USE
AC	0.92																
AES	**.41	0.76															
AI	**.49	**.55	0.92														
ATT	.09	*.15	.06	1.00													
BUSE	*.14	**.20	.10	*.14	1.00												
BEAU	**.17	**.54	**.23	**.19	*.13	1.00											
BI	**.39	**.41	**.39	.07	**.30	**.24	0.97										
GOOD	**.32	**.50	**.36	**.15	**.24	**.54	**.42	1.00									
HQ	**.26	**.64	**.41	**.15	**.15	**.61	**.37	**.62	0.78								
NEG	-.12	-.03	-.14	*.13	.03	-.03	*-.16	**-.18	-.02	0.68							
PD	**-.48	**-.23	**-.25	-.01	**-.16	-.11	**-.40	**-.35	**-.26	**.30	0.83						
PE	**.34	**.54	**.37	**.19	**.37	**.41	**.49	**.53	**.52	-.10	**-.35	0.91					
POS	**.20	**.41	**.32	**.21	**.26	**.34	**.27	**.30	**.36	**.18	-.16	**.54	0.68				
PQ	**.48	**.39	**.37	.09	**.23	**.32	**.43	**.52	**.45	**-.22	**-.51	**.42	**.25	0.73			
PUID	**.54	**.60	**.43	*.13	**.19	**.36	**.53	**.47	**.45	-.11	**-.52	**.56	**.31	**.60	0.93		
UC	**.55	**.42	**.63	.03	.09	*.14	**.46	**.34	**.32	**-.20	**-.29	**.38	**.26	**.41	**.47	0.92	
USE	**.15	**.26	*.13	*.13	**.81	**.17	**.33	**.27	**.18	.06	*-.15	**.40	**.27	**.22	**.22	.10	1.00

Note. Diagonal elements are square root of average variance extracted for each variable. Off-diagonal elements are correlation coefficients. AC: accessibility. AES: aesthetics. AI: adequacy of information. ATT: attribution. BUSE: baseline use frequency. BEAU: beauty. BI: behavioural intention. GOOD: goodness. HQ: hedonic quality. NEG: negative affect. PD: perceived disorientation. PE: perceived enjoyment. POS: positive affect. PQ: pragmatic quality. PUID: perceived user-interface design. UC: usefulness of content. USE: use frequency.

*$p < .05$. **$p < .01$.

Square roots of *AVE* values exceeded values of scale inter-correlations, supporting discriminant validity for each scale, according to the Fornell-Larcker criterion (Fornell & Larcker, 1981). According to Henseler et al. (2009), the loading of each reflective measurement item should be .70 or higher (50% or more variance shared by the item and the construct of which the item is an indicator). Note that several items in Table 4.13 had lower loadings than .70; however, these items were retained to facilitate comparability with previous studies. Descriptive statistics for each construct are presented in Table 4.15 (means and standard deviations were calculated using latent variable scores produced by SmartPLS).

Table 4.15. Descriptive statistics for each construct measuring aspects of interaction experience in Study 2.

Construct	Number of items	Mean	Standard deviation
Perceived aesthetics	8	4.80	1.09
Perceived disorientation	7	2.07	1.20
Perceived user-interface design	3	5.61	1.16
Usefulness of content	2	6.34	0.91
Adequacy of information	2	5.59	1.16
Accessibility	2	6.08	0.99
Pragmatic quality	4	5.49	1.04
Hedonic quality	4	4.76	1.09
Perceived enjoyment	3	4.96	1.24
Positive affect	10	3.77	1.16
Negative affect	10	1.94	0.92
Behavioural intention	2	6.26	1.18
Beauty	1	4.22	1.20
Goodness	1	5.44	1.32

4.5. Discussion

This chapter presented a selection of measures of aspects of users' interaction experience with news sites and Study 2, an on-line study that collected data on these

measures. A series of factor analyses and PLS analyses were carried out to establish the factor structure, internal consistency, construct validity and discriminant validity of the measures. *Classical aesthetics* and *expressive aesthetics* failed to reproduce as two distinct factors of aesthetic experience. However, statistical analysis supported the use of perceived aesthetics items as a unidimensional measure. One item was removed from each of the scales *usefulness of content* and *adequacy of information*, because of high cross-loadings. The remaining items loaded highly on the appropriate factor and produced low cross-loadings, indicating simple structure. Factor analyses of need-fulfilment subscales did not result in an interpretable structure; therefore, these scales were excluded from further analysis. In a general PLS measurement model, the confusing-structured (PQ2) item of *pragmatic quality* produced a low loading, but it was retained, based on psychometric considerations and to support comparability of results with previous studies. The scales *positive affect* and *negative affect* produced *AVE* values lower than .50 (.46 for both scales). However, the scales' construct validity was supported at the item level; therefore, they were retained for further analysis. The internal consistency, discriminant validity and construct validity of each scale was supported and the factor structure of the measures was confirmed. A structural model based on the data of Study 2 is presented in the next chapter.

Chapter 5

A structural model of interaction experience with news sites

Contents

5.1. Overview

The previous chapter introduced a measurement model of interaction experience with new sites, examining the factor structure, internal consistency, convergent validity and discriminant validity of measures collected in an online survey. This chapter, drawing on the same data set, presents the structural relationships between these measures. The chapter starts with the presentation of hypotheses, based on the literature from which the constructs of the study were adopted. Hypotheses 1 to 13 pertain to a model of interaction experience, based on the user-experience model (Hassenzahl, 2004). Hypotheses 14 to 23 pertain to a technology acceptance model (based on Davis, 1989), without the inclusion of *perceived ease of use* among the belief constructs, but augmented with *perceived enjoyment* (e.g., Cyr et al., 2006). Hypotheses 24 to 31, based on the work of van Schaik and Ling (2011), pertain to an integrated experience-acceptance model, including all technology-acceptance and interaction-experience variables, but not their antecedents. The different models are tested to examine the relationships between the variables of the measurement model presented in the previous chapter, based on hypotheses derived from previous research. Furthermore, testing different models is motivated by the idea that, depending on outcomes of interaction that are of interest (e.g., intention to use and overall product evaluations), different models are appropriate. Therefore, the choice of outcome variables would lead to the choice of a particular model. The presentation of the hypotheses is followed, first, by the testing of an interaction experience model, second, by the testing of a technology acceptance model, and third, by the testing of an experience-acceptance model. Finally, a model of interaction experience with news sites is presented, drawing on all the variables and

significant paths of the previous models, based on the components of user-experience model (Thüring & Mahlke, 2007; see Chapter 2).

5.2. Hypotheses

5.2.1. Hypotheses related to an interaction-experience model

The hypotheses related to the proposed interaction-experience model are presented in Figure 5.1. Based on Hassenzahl's user-experience model (2003, 2004; see Chapter 2) and empirical studies that confirmed it in the context of the use of Web sites (van Schaik & Ling, 2008, 2011), the following hypotheses are proposed:

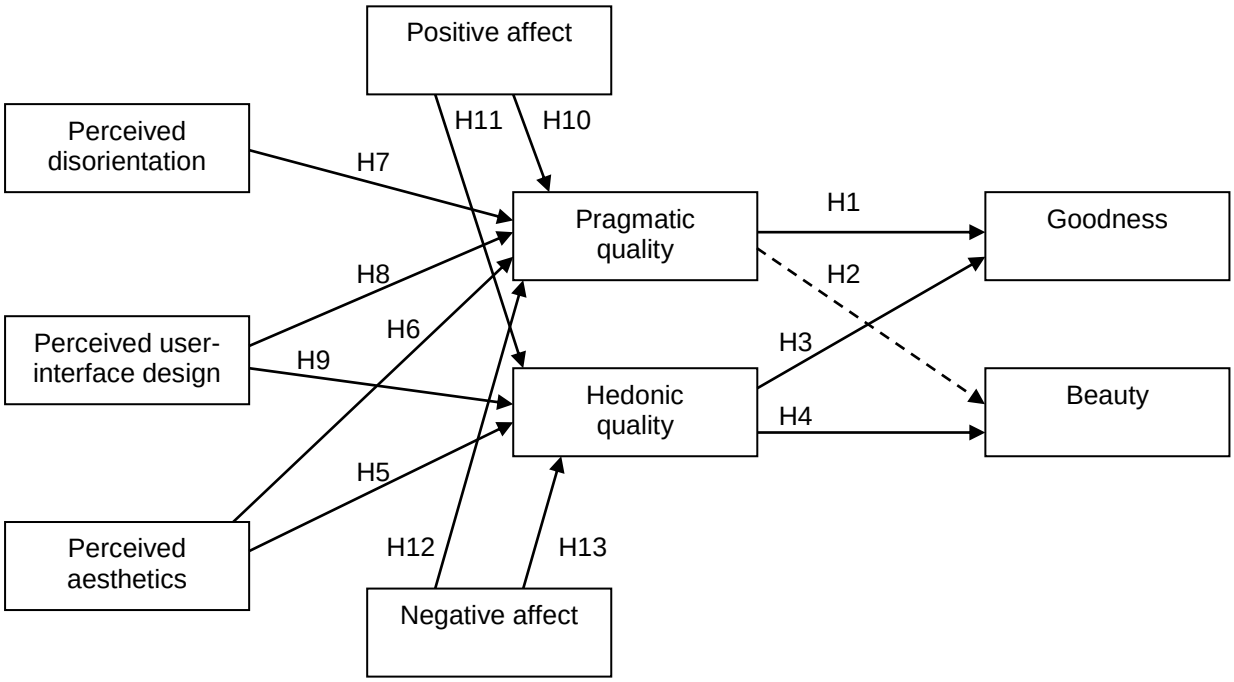


Figure 5.1. Hypotheses related to an interaction-experience model.

Hypothesis 1: *pragmatic quality* is a determinant of *goodness*.

Hypothesis 2: *pragmatic quality* is not a determinant of *beauty*.

Hypothesis 3: *hedonic quality* is a determinant of *goodness*.

Hypothesis 4: *hedonic quality* is a determinant of *beauty*.

Visual aesthetics is considered an important non-instrumental product characteristic in the interaction-experience literature (see Lavie & Tractinsky, 2004; Hassenzahl & Tractinsky, 2006; Hartmann et al., 2008). Theoretically, *expressive aesthetics* is expected to be an antecedent of *hedonic quality*, because the latter is conceptualised as the pleasure-producing qualities of a particular artefact, and, as such, is a determinant of *beauty* perceptions. *Classical aesthetics*, on the other hand, emphasises clearly and orderly design and is also expected to be connected to *pragmatic quality*, as it was found to be related to evaluations of usability (e.g., Lavie & Tractinsky, 2004; van Schaik & Ling, 2009). The analysis of interaction-experience measures in the exploratory study confirmed these connections between the classical and expressive aesthetics dimensions and product attributes (see Chapter 3). However, the two aesthetics dimensions did not reproduce in the measurement model and a composite of the two scales were identified as psychometrically justified solution to measure aesthetics. Therefore, in order to explore the relationships between perceived aesthetics and perceived product attributes, the following hypotheses are proposed:

Hypothesis 5: *perceived aesthetics* is an antecedent of *hedonic quality*.

Hypothesis 6: *perceived aesthetics* is an antecedent of *pragmatic quality.*

Ahuja and Webster (2001) found that disorientation and ease of use are distinct, but strongly and negatively related constructs. In an experiment involving information retrieval tasks from a Web site, Van Schaik and Ling (2003) confirmed that disorientation and ease of use are different constructs, and by manipulating

orientation support, they found that *perceived disorientation* is more sensitive measure than *perceived ease of use*. Because *pragmatic quality* is essentially the user-perceived usability of a particular artefact (Hassenzahl, 2004), it is reasonable to assume that disorientation effects the pragmatic quality perceptions of a particular artefact. Therefore, the following hypothesis is proposed:

Hypothesis 7: *perceived disorientation* is an antecedent of *pragmatic quality*.

User-interface design concerns the presentation of the interface of a particular technology. The *perceived user-interface design* measure adopted for the current study (based on Cho et al., 2009) emphasises the layout of a Web site, that is, whether the functional and graphic elements are presented appropriately. Presentation is part of product features in Hassenzahl's model of interaction experience (Hassenzahl, 2004), and therefore it is expected to influence the perception of product attributes. The layout of functional and graphic elements (e.g., textual and multimedia content and links) on the pages of a particular news site may influence perceived usability, and at the same time, it is fundamentally connected to the appearance of the site. To examine the relationships between *perceived user-interface design* and perceived product attributes, the following hypotheses are proposed:

Hypothesis 8: *perceived user-interface design* is an antecedent of *pragmatic quality*.

Hypothesis 9: *perceived user-interface design* is an antecedent of *hedonic quality*.

Hassenzahl and colleagues (2010) found that *hedonic quality* is positively related to *positive affect* ($r = .46$, $p < .001$), and found support that it remains a significant predictor of *hedonic quality* after controlling for the effect of need fulfilment. Based

on mediation and moderation analyses, the authors theorised that *positive affect* is an outcome of need fulfilment and "a legitimate predictor of hedonic quality" (p. 361). Similarly, they found that *positive affect* is positively correlated with *pragmatic quality* (r = .28, p < .001) and it was a significant predictor of *pragmatic quality*. In the measurement model, the scale inter-correlation between *hedonic quality* and *positive affect* was medium (r = .35 p < .01), and the scale inter-correlation between *pragmatic quality* and *positive affect* was small (r = .22, p < .01). (See Chapter 4, Table 4.14 for inter-scale correlations of the measurement model.) With regards to *negative affect*, an opposite effect is expected on product attributes. It is expected that *negative affect* experienced during the interaction results in lowered ratings of *pragmatic quality* and *hedonic quality*. To investigate the connections between affect and product attributes, the following hypotheses are proposed:

Hypothesis 10: *positive affect* is an antecedent of *pragmatic quality*.

Hypothesis 11: *positive affect* is an antecedent of *hedonic quality*.

Hypothesis 12: *negative affect* is an antecedent of *pragmatic quality*.

Hypothesis 13: *negative affect* is an antecedent of *hedonic quality*.

5.2.2. Hypotheses related to a technology acceptance model

The hypotheses related to the proposed technology acceptance model are presented in Figure 5.2. *Perceived usefulness* in this study is characterised as the perceived usefulness of the content presented by the news portal a particular participant has been using. Therefore, *usefulness of content* is used as a proxy of *perceived usefulness* in this analysis. Based on the technology acceptance model (e.g., Davis,

1989; Venkatesh & Davis, 2000; see Chapter 2), the following hypothesis is proposed:

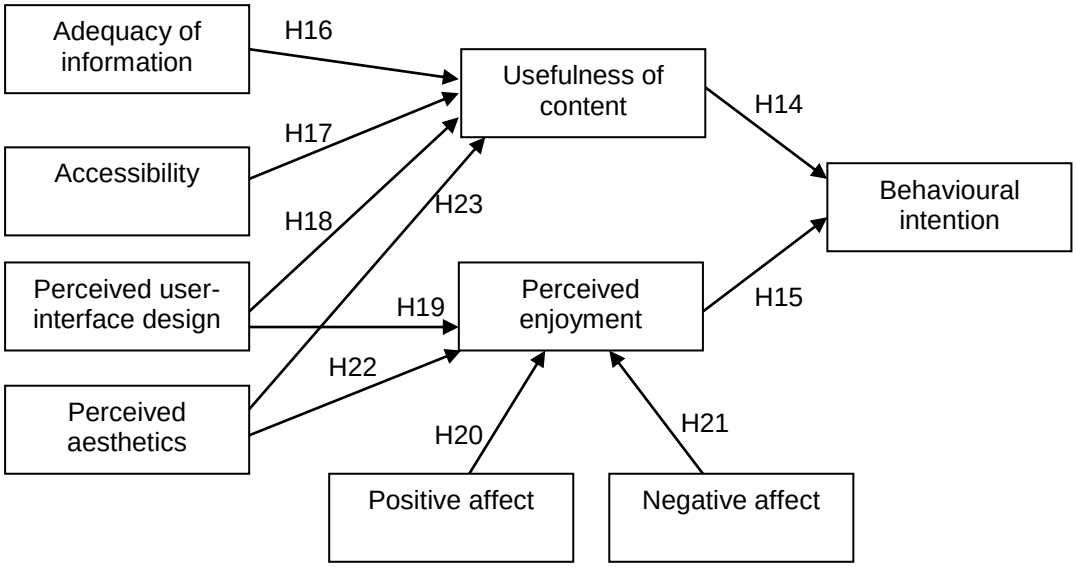


Figure 5.2. Hypotheses related to a technology acceptance model.

Hypothesis 14: *usefulness of content* is a determinant of *behavioural intention*.

P*erceived enjoyment* in Chapter 2 was introduced as an antecedent of *perceived ease of use*, in order to demonstrate how the belief constructs of the original technology acceptance model was augmented with antecedents in the literature. However, research confirmed that *perceived enjoyment* is a direct determinant of *behavioural intention* (Cyr et al., 2006, 2007; van Schaik & Ling, 2011). Furthermore, *perceived ease of use* is not included in the present research, consequently, the mediation of *perceived enjoyment* through *perceived ease of use* (Sun & Zhang, 2008) cannot be tested and a direct effect of *perceived enjoyment* on *behavioural intention* can be expected. Therefore, the following hypothesis is proposed:

Hypothesis 15: *perceived enjoyment* is a determinant of *behavioural intention*.

The measures of *usefulness of content*, *adequacy of information* and *accessibility* were adopted from the same instrument measuring user-perceived service quality of information-presenting Web portals (Yang et al., 2005). In the development and validation of the original instrument, *adequacy of information*, along with *usefulness of content*, was conceptually presented as a determinant of information quality, whereas *accessibility* was a determinant of service quality. In turn, both information- and service quality were determinants of acceptance of technology. *Usefulness of content*, however, is used here as a proxy of *perceived usefulness*. In the adoption of the technology acceptance model for the domain of news sites (Chen & Corkindale, 2008; see Chapter 2), *perceived core service quality*, the quality and presentation of content, is described as an antecedent of *perceived usefulness*. It is proposed that *adequacy of information*, as a measure of information quality, and *accessibility*, as a measure of service quality, are antecedents of *usefulness of content*, which, in turn, is a determinant of acceptance of technology. Therefore, the following hypotheses are proposed:

Hypothesis 16: *adequacy of information* is an antecedent of *usefulness of content*.

Hypothesis 17*: accessibility* is an antecedent of *usefulness of content*.

In an adaptation of the technology acceptance model to Web technologies (Cho et al., 2009; see Chapter 2), *perceived user-interface design* is an antecedent of *perceived usefulness* and *perceived ease of use*. In the model, the effect of *perceived user-interface design* on *perceived usefulness* is mediated through *perceived functionality*. Because there is no measure of *perceived functionality* involved in the current study, a direct effect *of perceived user-interface design* is expected on *usefulness of content*. In the absence of a measure of *perceived ease*

of use in the current study, *perceived user-interface design* is expected to have a direct effect on *perceived enjoyment*. Therefore, the following hypotheses are proposed:

Hypothesis 18: *perceived user-interface design* is an antecedent of *usefulness of content*.

Hypothesis 19*: perceived user-interface design* is an antecedent of *perceived enjoyment*.

Perceived enjoyment is, by definition, an intrinsic motivation variable (Sun & Zhang, 2008; see Chapter 2) that changes over time and across artefacts. In effect, it may be considered as a state-affect variable, and therefore it is expected to be connected to affective reactions measured in relation to the interaction with a particular artefact. Therefore, the following hypotheses are proposed:

Hypothesis 20: *positive affect* is an antecedent of *perceived enjoyment*.

Hypothesis 21: *negative affect* is an antecedent of *perceived enjoyment*.

Regarding the role of aesthetics in technology acceptance, two studies (van der Heijden, 2004; Cyr et al., 2006) have independently established that *perceived aesthetics* is an antecedent of *perceived enjoyment*, as well as of *perceived usefulness* and *perceived ease of use*. Therefore, the following hypotheses are proposed:

Hypothesis 22: *perceived aesthetics* is an antecedent of *perceived enjoyment*.

Hypothesis 23: *perceived aesthetics* is an antecedent of *usefulness of content*.

5.2.3. Hypotheses related to an integrated experience-acceptance model

The hypotheses related to the proposed integrated experience-acceptance model are presented in Figure 5.3. In an integrated model of interaction experience for information retrieval in a Web-based encyclopaedia, van Schaik and Ling (2011) found that perceptions of product attributes (*pragmatic quality* and *hedonic quality*) are independent determinants of technology-acceptance constructs (*perceived usefulness*, *perceived ease of use* and *perceived enjoyment*), but product evaluations (beauty and goodness) are not independent determinants of *intention to use*. Therefore, the following hypotheses are proposed:

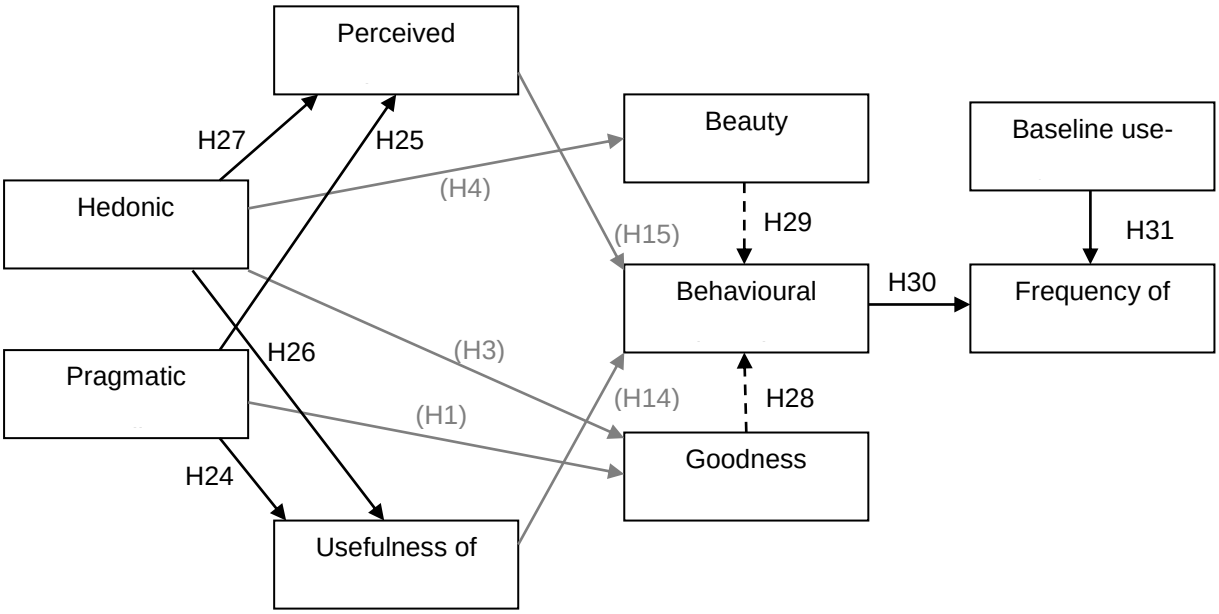


Figure 5.3. Hypotheses related to an integrated experience-acceptance model.

Hypothesis 24: *pragmatic quality* is a determinant of *usefulness of content*.

Hypothesis 25: *pragmatic quality* is a determinant of *perceived enjoyment*.

Hypothesis 26: *hedonic quality* is a determinant of *usefulness of content*.

Hypothesis 27: *hedonic quality* is a determinant of *perceived enjoyment*.

Hypothesis 28: *goodness* is not a determinant of *behavioural intention*.

Hypothesis 29: *beauty* is not a determinant of *behavioural intention*.

Based on the technology acceptance model (Davis, 1989), it is expected that *behavioural intention* is positively related to use behaviour. Use behaviour in the study was assessed by how frequently a particular participant normally accessed the news site that he/she had been using during the study (*frequency of use*). However, the participants of Study 2 were all in the post-adoption stage (Magni et al., 2010), in other words, they have already adopted the news sites they chose to use during the study. Because of this, their actual use-frequency of the particular news sites is expected to be less influenced by *behavioural intention*. Participants were asked to indicate how frequently they accessed news sites in general (*baseline use-frequency*), as well as the frequency of use of the particular news site they had been using during the study. The frequency of use of news sites in general is expected to be positively related to the frequency of use of a particular news site. Therefore, the following hypotheses are proposed:

Hypothesis 30: *behavioural intention* is a determinant of *frequency of use*.

Hypothesis 31: *baseline use-frequency* is a determinant of *frequency of use*.

Note that the intention to use (or continue using) a particular news site in the future is not necessarily informative for predicting the frequency of use of that site. The frequency of use of news sites in general, however, is expected to be more strongly related to the frequency of use of particular sites. Therefore, Hypothesis 30 is proposed to test if behavioural intention still remains a significant predictor of use behaviour after controlling for baseline use.

5.3. Analyses

5.3.1. Interaction experience model

Based on Hypotheses 1-13 (see Figure 5.1), the following variables were entered the analysis of an interaction experience model with news sites: *perceived disorientation*, *perceived user-interface design*, *perceived aesthetics*, the affect dimensions of *positive affect* and *negative affect*, the perceived product characteristic measures of *pragmatic quality* and *hedonic quality*, and the evaluative judgements of *beauty* and *goodness*. The results of the hypothesis tests are summarised in Table 5.1.

Table 5.1. Summary of the path analysis of interaction-experience variables.

	Predictor variable	Target variable	β	t^a	sr^2	Unique effect size (f^2)[b]
H_1	Pragmatic quality	Goodness	.31	***5.93	.08	.14
H_2	Pragmatic quality	Beauty	.06	1.11	.00	.00
H_3	Hedonic quality	Goodness	.48	***9.19	.18	.33
H_4	Hedonic quality	Beauty	.59	***12.72	.28	.45
H_5	Perceived aesthetics	Hedonic quality	.53	***9.06	.17	.29
H_6	Perceived aesthetics	Pragmatic quality	.04	0.74	.00	.00
H_7	Perceived disorientation	Pragmatic quality	-.24	***3.50	.04	.07
H_8	Perceived user-interface design	Pragmatic quality	.41	***5.66	.08	.15
H_9	Perceived user-interface design	Hedonic quality	.10	1.75	.01	.01
H_{10}	Positive affect	Pragmatic quality	.09	1.85	.01	.01
H_{11}	Positive affect	Hedonic quality	.13	*2.36	.01	.02
H_{12}	Negative affect	Pragmatic quality	-.12	*2.30	.01	.02
H_{13}	Negative affect	Hedonic quality	-.01	0.26	.00	.00

[a]Bootstrap, $N = 5000$.

[b]f^2: 0.02 - 0.14 small, 0.15 - 0.34 medium and 0.35 - large (Cohen, 1988).

*$p < .05$. ***$p < .001$.

Because Hypothesis 2 was supported, the path leading from *pragmatic quality* to *beauty* is omitted in the following analyses. Hypotheses 6, 9, 10 and 13 were not supported and these paths were also removed from the model. Although Hypotheses 11 and 12 were supported, *positive affect* and *negative affect* did not contribute substantially to the prediction of *hedonic quality* and *pragmatic quality*, respectively. Because of their small effect sizes, the affect dimensions were also

removed from the interaction-experience model. However, the pattern of connection of affective reactions to product attributes is noteworthy: *positive affect* was positively and significantly related to *hedonic quality*, and *negative affect* was negatively and significantly related to *pragmatic quality*.

After removing *positive affect* and *negative affect* from the model, the path leading from *perceived aesthetics* to *pragmatic quality* (Hypothesis 6) and the path from *perceived user-interface design* to *hedonic quality* (Hypothesis 9) were tested again, to examine if the presence of affect variables suppressed these paths. The connection between *perceived aesthetics* and *pragmatic quality* remained non-significant (β= .07, t = 1.29, *ns*). The connection between *perceived user-interface design* and *hedonic quality* approached significance, but remained above the threshold of 5% alpha (β = .12, t = 1.93, *ns*). Therefore, both paths were excluded from the model.

Following the exclusion of the non-significant paths and predictors with low impact on endogenous variables, mediation analyses were carried out to test if the effects of *perceived disorientation*, *perceived user-interface design* and *perceived aesthetics* on quality judgements are mediated through perceived product attributes (*pragmatic quality* and *hedonic quality*)[17]. The direct effect of *perceived disorientation* on *goodness* (β = -.36, t = 6.38, p < .001)[18] became non-significant with the introduction of *pragmatic quality* as a mediator (β = -.11, t = 1.72, *ns*), demonstrating full

[17] PLS models for testing mediation effects included a predictor variable, a target variable and a mediator variable. Full mediation is a case where the inclusion of a significant mediator into a model changes the path from the predictor variable to the target variable to non-significant; partial mediation is a case where the direct effect of the predictor becomes smaller, but remains statistically significant, as a result of including a significant mediator variable into a model (see Chin, 2010).

[18] A bootstrap sample of N = 5000 was used for all mediation analyses.

mediation. *Perceived disorientation* alone accounted for 26% of variance in *pragmatic quality*.

The direct effect of *perceived user-interface design* on *goodness* (β = .47, t = 7.91, p < .001) remained significant with the inclusion of *pragmatic quality* as a mediator (β = .23, t = 2.91, p < .01). However, with the inclusion of all other variables in the model, the direct effect became non-significant (β = .11, t = 1.47, *ns*). Therefore, the direct path leading from *perceived user-interface design* to *goodness* was not included in the model. *Perceived user-interface design* alone accounted for 37% variance in *pragmatic quality*.

The direct path between *perceived aesthetics* and *beauty* (β = .55, t = 10.40, p < .001) remained significant with the introduction of *hedonic quality* as a mediator (β = .25, t = 3.09, p < .01). The direct path remained significant after the introduction of all other variables (β = .25, t = 3.22, p < .01); therefore, it was retained in the model. The model parameters of the interaction experience model are presented in Table 5.2. The interaction experience path model is presented in Figure 5.4.

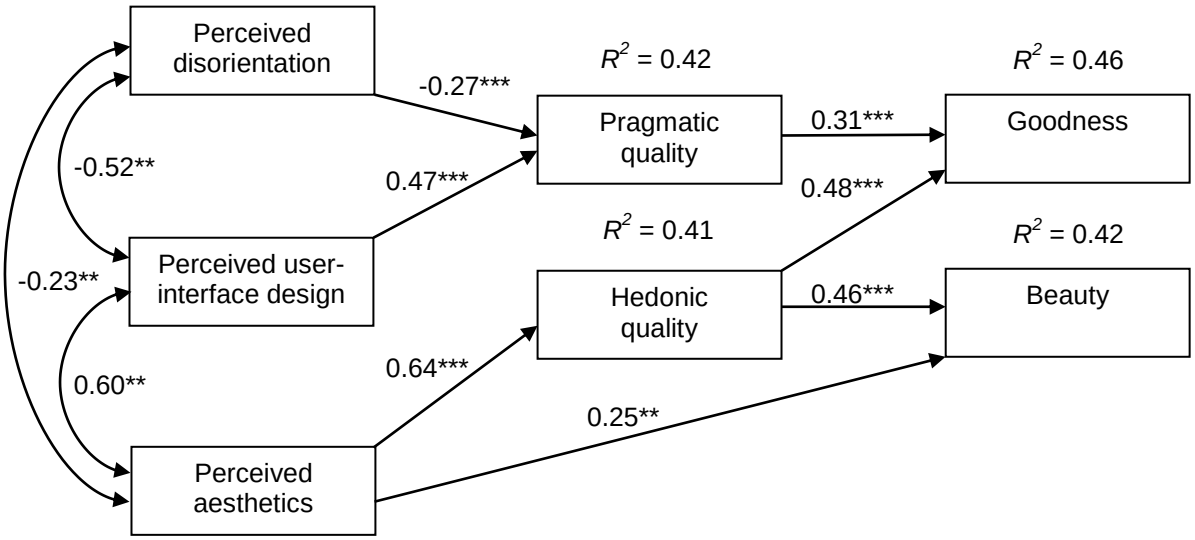
Table 5.2. Model parameters of the interaction experience model.

Predictor variable	Target variable	β	t[a]	sr^2	Unique effect size (f^2)[b]
Pragmatic quality	Goodness	.31	***6.14	.08	.14
Hedonic quality	Goodness	.48	***9.06	.18	.33
Hedonic quality	Beauty	.46	***5.85	.12	.21
Perceived disorientation	Pragmatic quality	-.27	***4.03	.05	.09
Perceived user-interface design	Pragmatic quality	.47	***7.24	.16	.28
Perceived aesthetics	Hedonic quality	.64	***14.55	.41	.71
Perceived aesthetics	Beauty	.25	**3.20	.04	.06

[a]Bootstrap, N = 5000.

[b]f^2: 0.02 - 0.14 small, 0.15 - 0.34 medium and 0.35 - large.

**p < .01. ***p < .001.



**p < .01. ***p < .001.

Figure 5.4. An interaction experience model of news Web sites.

5.3.2. Technology acceptance model

Based on Hypotheses 14-23 (see Figure 5.2), the following variables were entered the analysis of a technology acceptance model of interaction with news sites: *adequacy of information*, *accessibility*, *perceived user-interface design*, *perceived aesthetics*, the affect dimensions of *positive affect* and *negative affect*, the belief constructs of *usefulness of content* (as a proxy of *perceived usefulness*) and *perceived enjoyment*, and *behavioural intention*. The results of the hypothesis tests are summarised in Table 5.3.

Table 5.3. Summary of the path analysis of technology-acceptance variables.

	Predictor variable	Target variable	β	t^a	sr^2	Unique effect size (f^2)[b]
H_{14}	Usefulness of content	Behavioural intention	.32	***4.60	.09	.14
H_{15}	Perceived enjoyment	Behavioural intention	.38	***5.88	.12	.18
H_{16}	Adequacy of information	Usefulness of content	.46	***6.87	.13	.24
H_{17}	Accessibility	Usefulness of content	.25	**2.61	.04	.08
H_{18}	Perceived user-interface design	Usefulness of content	.13	1.66	.01	.02
H_{19}	Perceived user-interface design	Perceived enjoyment	.31	***5.87	.06	.12
H_{20}	Positive affect	Perceived enjoyment	.39	***10.23	.13	.26
H_{21}	Negative affect	Perceived enjoyment	-.15	*2.55	.02	.04
H_{22}	Perceived aesthetics	Perceived enjoyment	.19	***3.44	.02	.04
H_{23}	Perceived aesthetics	Usefulness of content	-.01	0.17	.00	.00

[a]Bootstrap, $N = 5000$.

[b]f^2: 0.02 - 0.14 small, 0.15 - 0.34 medium and 0.35 - large.

**$p < .01$. ***$p < .001$.

Mediation analysis showed that effects of both affect dimensions on *behavioural intention* were fully mediated through *perceived enjoyment*. The direct path leading from *positive affect* to *behavioural intention* (β = .37, $t = 8.96$, $p < .001$) became non-significant with the introduction *of perceived enjoyment* as a mediator in the model (β = .00, $t = 0.05$, *ns*). Similarly, the direct effect of *negative affect* (β = -.20, $t = 2.23$, $p = .05$) became non-significant with the introduction of *perceived enjoyment* (β = -.12, $t = 1.49$, *ns*). With *perceived user-interface design* and *perceived aesthetics* excluded, *positive affect* and *negative affect* accounted for the 34% of variance in *perceived enjoyment*.

The direct effect of *perceived aesthetics* on *behavioural intention* (β = .43, $t = 8.49$, $p < .001$) remained significant with the introduction of *perceived enjoyment* as a mediator (β = .21, $t = 3.48$, $p < .001$). However, in the complete model, with all other variables and paths included, the direct path lost significance (β = .02, $t = 0.24$, *ns*). When *perceived user-interface design* was excluded, the effect size of *perceived*

aesthetics on *perceived enjoyment* was medium (f^2 = .21). However, the inclusion of *perceived user-interface design* lowered this to small (f^2 = .05), suggesting that the two variables share a significant portion of variance. *Perceived aesthetics* was retained in the model, because the path leading from it to *perceived enjoyment* remained significant (β = .20, *t* = 3.66, *p* < .001).

Perceived user-interface design retained a significant portion of variance explained in *behavioural intention* after the introduction of *perceived enjoyment* as a mediator (β = .37, *t* = 5.73, *p* < .001), and this effect remained significant in the complete model. Therefore, both the direct and indirect paths of *perceived user-interface design* leading to *behavioural intention* were included in the model.

The direct effect of *adequacy of information* (β = .15, *t* = 2.17, *p* < .05) on *behavioural intention* became non-significant with the introduction of *usefulness of content* as a mediator (β = .15, *t* = 1.67, *ns*). The direct effect of *accessibility* on *behavioural intention* (β = .40, *t* = 6.43, *p* < .01) remained significant after the introduction of *usefulness of content* (β = .20, *t* = 2.45, *p* < .05), suggesting partial mediation. However, in the complete model, the direct path lost significance (β = .04, *t* = 0.52, *ns*). Together, the two constructs accounted for the 48% of variance in *usefulness of content*.

The paths of *perceived aesthetics* and *perceived user-interface design* to *usefulness of content* were removed from the model because of their lack of significance. The model parameters of the technology acceptance model are presented in Table 5.4. The technology acceptance path model is presented in Figure 5.5.

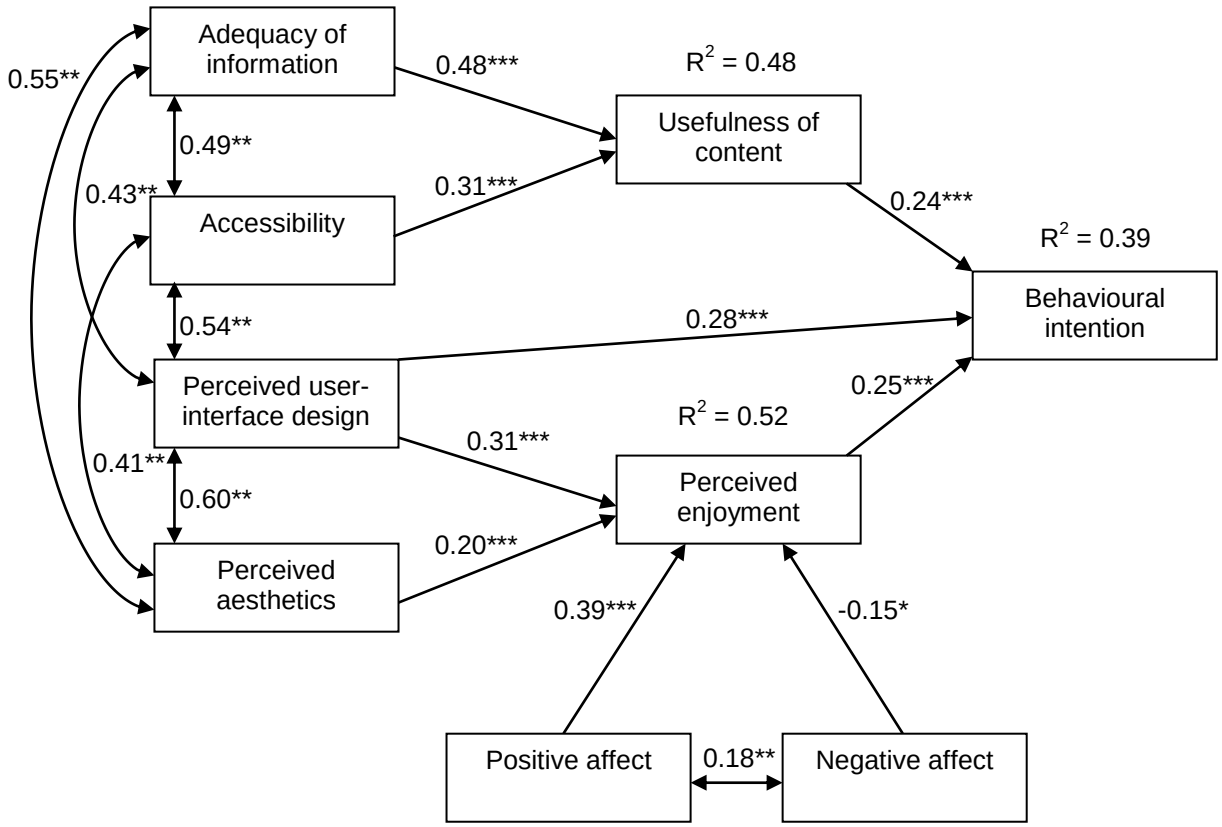
Table 5.4. Model parameters of the technology acceptance model.

Predictor variable	Target variable	β	t^a	sr^2	Unique effect size (f^2)[b]
Usefulness of content	Behavioural intention	.24	***3.37	.05	.07
Perceived user-interface design	Behavioural intention	.28	***4.26	.05	.08
Perceived enjoyment	Behavioural intention	.25	***4.04	.04	.07
Adequacy of information	Usefulness of content	.48	***6.87	.17	.33
Accessibility	Usefulness of content	.31	***4.07	.07	.14
Perceived user-interface design	Perceived enjoyment	.31	***5.79	.06	.13
Positive affect	Perceived enjoyment	.39	***10.17	.12	.26
Negative affect	Perceived enjoyment	-.15	*2.51	.02	.05
Perceived aesthetics	Perceived enjoyment	.20	***3.57	.02	.05

[a]Bootstrap, N = 5000.

[b]f^2: 0.02 - 0.14 small, 0.15 - 0.34 medium and 0.35 - large.

*p < .05. ***p < .001.



*p < .05, ***p < .001

Figure 5.5. A technology acceptance model of news Web sites.

5.3.3. Integrated experience-acceptance model

In order to test an integrated experience-acceptance model, the following measures were entered the analysis: perceived product attributes (*pragmatic quality* and *hedonic quality*), belief constructs of the technology acceptance model (*usefulness of content* and *perceived enjoyment*), product judgements (*goodness* and *beauty*), *behavioural intention*, *frequency of use* and *baseline use-frequency*. The results of the hypothesis tests are summarised in Table 5.5. In addition to the paths specified in Hypotheses 24-31 (see Figure 5.3), the paths leading from *pragmatic quality* to *goodness*, from *hedonic quality* to *goodness* and *beauty* and from *usefulness of content* and *perceived enjoyment* to *behavioural intention* were also included in the analysis, because they were supported in the test of the interaction experience and technology acceptance models. All of these paths retained significance and are not included in Table 5.5.

Table 5.5. Summary of hypothesis tests of an integrated experience-acceptance model.

	Predictor variable	Target variable	β	t[a]	sr^2	Unique effect size (f^2)[b]
H_{24}	Pragmatic quality	Usefulness of content	.34	***5.30	.09	.11
H_{25}	Pragmatic quality	Perceived enjoyment	.24	***4.60	.05	.07
H_{26}	Hedonic quality	Usefulness of content	.17	*2.45	.02	.03
H_{27}	Hedonic quality	Perceived enjoyment	.42	***6.46	.14	.20
H_{28}	Goodness	Behavioural intention	.17	**2.58	.02	.02
H_{29}	Beauty	Behavioural intention	-.01	0.10	.00	.00
H_{30}	Behavioural intention	Frequency of use	.10	*2.29	.01	.02
H_{31}	Baseline use-frequency	Frequency of use	.78	***21.49	.55	1.64[c]

[a]Bootstrap, N = 5000.

[b]f^2: 0.02 - 0.14 small, 0.15 - 0.34 medium and 0.35 - large.

[c]An effect size coefficient of f^2 may exceed 1.00, in particular when the R^2 of a particular dependent variable is close to 1.00 (see Chin, 2010 for the formula of calculating f^2).

*p < .05. **p < .01. ***p < .001.

Except for Hypothesis 28, all hypotheses were supported. *Goodness* was a significant predictor of *behavioural intention* (possibly because *perceived ease of use* was missing from the model); however, its effect size was small and it accounted for only 2% of unique variance in *behavioural intention*. Perceived product attributes were significant predictors of technology acceptance constructs. *Pragmatic quality* was most strongly related to *usefulness of content*, whereas *hedonic quality* was more strongly related to *perceived enjoyment*. *Behavioural intention* alone accounted for 11% of the variance in *frequency of use*, but it only accounted for 1% of unique variance in *frequency of use* with *baseline use-frequency* included in the model. This suggests that in the post-adoption stage, the power of *behavioural intention* (the intention to use a particular artefact in the future) to predict the *frequency of use* of a particular new site is diminished; however, the relationship between the two variables remains significant.

The model parameters were recalculated after the removal of the non-significant path between *beauty* and *behavioural intention*. The model parameters of the integrated experience-acceptance model are presented in Table 5.6. The integrated experience-acceptance path model is presented in Figure 5.6.

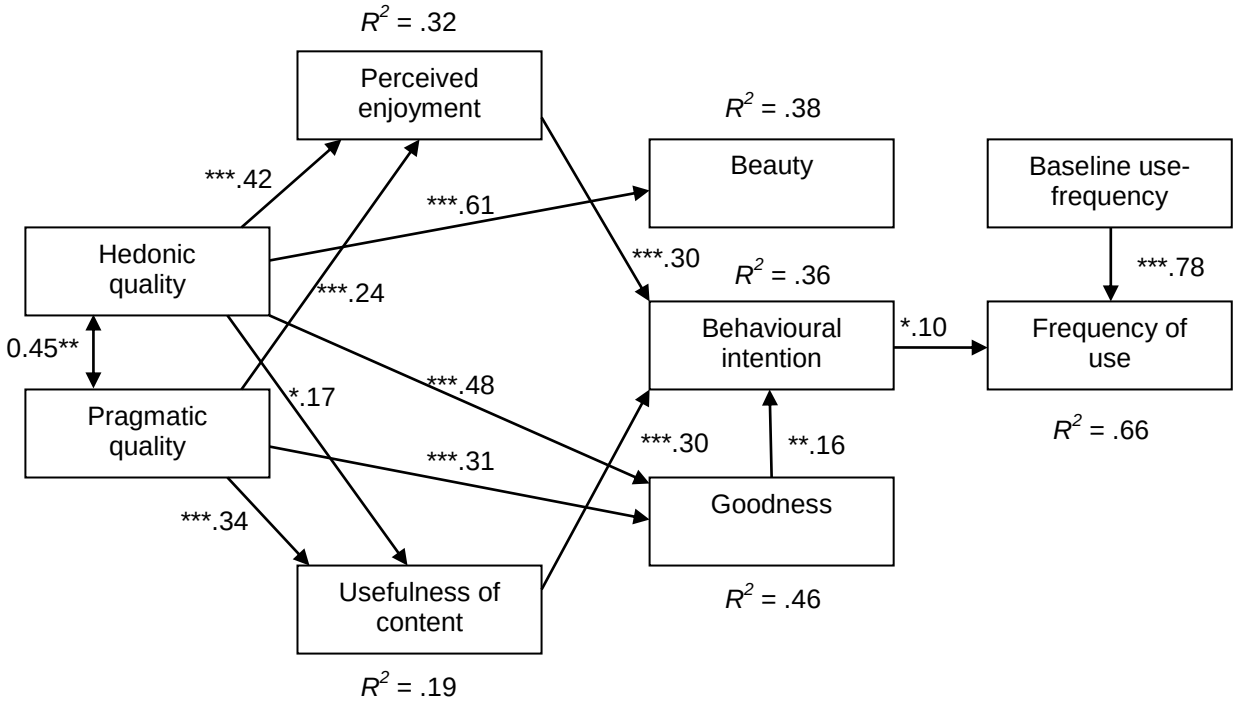
Table 5.6. Model parameters of the integrated experience-acceptance model.

Predictor variable	Target variable	β	t^a	sr^2	Unique effect size (f^2)[b]
Pragmatic quality	Goodness	.31	***5.93	.08	.14
Hedonic quality	Beauty	.61	***14.22	.38	.61
Hedonic quality	Goodness	.48	***9.23	.18	.34
Usefulness of content	Behavioural intention	.30	***4.12	.07	.11
Perceived enjoyment	Behavioural intention	.30	***5.00	.06	.09
Pragmatic quality	Usefulness of content	.34	***5.44	.09	.11
Pragmatic quality	Perceived enjoyment	.24	***4.59	.05	.07
Hedonic quality	Usefulness of content	.17	*2.44	.02	.03
Hedonic quality	Perceived enjoyment	.42	***6.37	.14	.20
Goodness	Behavioural intention	.16	**2.71	.02	.03
Behavioural intention	Frequency of use	.10	*2.35	.01	.02
Baseline use-frequency	Frequency of use	.78	***21.40	.55	1.64

[a]Bootstrap, N = 5000.

[b]f^2: 0.02 - 0.14 small, 0.15 - 0.34 medium and 0.35 - large.

*p < .05. **p < .01. ***p < .001.



*p < .05. **p < .01. ***p < .001.

Figure 5.6. An integrated experience-acceptance model of news Web sites.

5.3.4. A model of interaction experience with news sites

Chapter 2 concluded that the most complete conceptual model of interaction experience in the relevant human-computer interaction literature, in terms of the range of variables involved, is the components of user-experience model (Thüring & Mahlke, 2007). The model is composed of three groups of variables: interaction characteristics, components of user-experience and variables pertaining to the appraisal of a particular interactive system. Interaction characteristics are a set of artefact, person and task/context characteristics. Components of user-experience are perceptions of instrumental qualities, perceptions of non-instrumental qualities and emotional responses. Appraisal of the system may be considered as the outcome of the interaction and consist of measures such as the overall judgements of the system and use behaviour (see Chapter 2 and Figure 2.9. for a more detailed description). Here, components of user-experience are referred to instead as components of interaction experience, and appraisal of the system is referred to as interaction outcomes, because system appraisal, conceptualised as overall quality judgements of *beauty* and *goodness*, can be considered as one aspect of outcomes of interaction.

Based on the analyses of the interaction experience, technology acceptance and integrated experience-acceptance models of interaction with news sites, the variables in the study were arranged in terms of the components of user-experience model. The aim was to form an integrated model of interaction experience with news sites, which also incorporates variables of technology acceptance, but goes beyond the mere merging of the models of interaction experience and technology acceptance. The model of components of user-experience was selected to serve as a basis for the construction of the model, because it adheres to the person-artefact-

task framework, and the description of its particular components are flexible enough to provide the freedom to select variables for conceptualising measurement.

Regarding interaction characteristics, the first group of variables in the model, the current study only contains variables measuring artefact characteristics. These measures are: *adequacy of information*, *accessibility*, *perceived disorientation*, *perceived user-interface design* and *perceived aesthetics*. The connections of these measures with perceived product characteristics and with technology acceptance constructs were established during the hypothesis tests of the interaction experience and technology acceptance models in this chapter. The direct effect of *perceived aesthetics* on *beauty* in the interaction experience model and of *perceived user-interface design* on *behavioural intention* in the technology acceptance model, however, are omitted here for the sake of a simpler model-structure, and the variables of interaction outcomes are only predicted from variables of components of interaction experience.

Components of interaction experience are divided into three groups of constructs: perception of instrumental qualities, perception of non-instrumental qualities and emotional responses. In the current model, *usefulness of content* and *pragmatic quality* are regarded as measures of perception of instrumental qualities, and *perceived enjoyment* and *hedonic quality* as measures of perception of non-instrumental qualities. Emotional responses in the current model are characterised as *positive affect and negative affect* experienced during the interaction. Appraisal of the system in the current model is characterised with the outcome measures of the interaction experience and technology acceptance models: *behavioural intention*, and the evaluative judgement items of *beauty* and *goodness*. *Frequency of use* is not

used here, because in the integrated experience-acceptance model it was found to be more strongly related to *baseline use-frequency*, rather than to other endogenous variables of the model.

In the components of user-experience model, perceived product qualities are influenced by interaction characteristics; emotional responses are described as being a consequence of the perception of instrumental and non-instrumental qualities of a particular system and, in turn, they serve as determinants of the appraisal of the system. However, it may as well be that perceptions of artefact characteristics trigger an affective reaction, which then influences the perceptions of instrumental and non-instrumental product qualities, and serves as a direct determinant of system appraisal. In fact, affective reactions, like first-impression perceptions of aesthetics (Lindgaard et al, 2006), can form instantaneously and automatically, without cognitive processing (Zajonc, 1980), and thus are likely to form at an early stage of interaction with interactive technologies. However, cognitive appraisals, as another possible measure of emotional responses, are more cognitive in nature than affective reactions (Scherer, 2001) and might be considered as a consequence of perception of product qualities. The affective reactions measured in the current study can be considered as more closely related to Norman's (2005) 'visceral level'(Desmet & Hekkert, 2007), whereas more complex emotional responses arising from appraisal processes can be considered as more closely related to the 'behavioural' and 'reflective' levels. Therefore, in the current study, the path from interaction characteristics through affective reactions to perceived product qualities is preferred rather than the reverse relationship between affect and perceptions of product quality, proposed by Thüring and Mahlke (2007).

Furthermore, in the technology acceptance model in the current chapter, affective reactions were used to predict *perceived enjoyment*, and a full mediation of the effects of the two affect dimensions on *behavioural intention* was established. In the hypothesis testing of the interaction experience model, *positive affect* was found to significantly predict *hedonic quality* and *negative affect* to predict *pragmatic quality*, although these paths were omitted because of their low effect sizes. The paths proposed in the hypothesis tests of the interaction experience model between affect dimensions and perceived product attributes are re-introduced in the current model to examine the pattern of relationship between affective reactions and perceptions of instrumental and non-instrumental qualities.

The paths from the affect dimensions to *usefulness of content* are introduced in the model as an interaction measure. The paths leading from *negative affect* and *positive affect* to *goodness* and *beauty* evaluations are introduced to test if affective responses exert a direct effect on quality judgements, over and above their indirect effects through technology acceptance constructs and product attributes. Although in the original components of user-experience model, interaction characteristics are not used to predict emotional responses, the connections between measures of artefact characteristics and affect dimensions are tested here to see if affect experienced during interaction can be connected to designable product characteristics. As Hassenzahl (2006) points out, designers of interactive products cannot exert a high level of control over emotional responses in a particular design, but they can design to create the possibility of an experience to occur during future interactions. It is therefore useful to identify connections between artefact characteristics and emotional responses to aid designers. An outline of the model of interaction experience with news sites is presented in Figure 5.7.

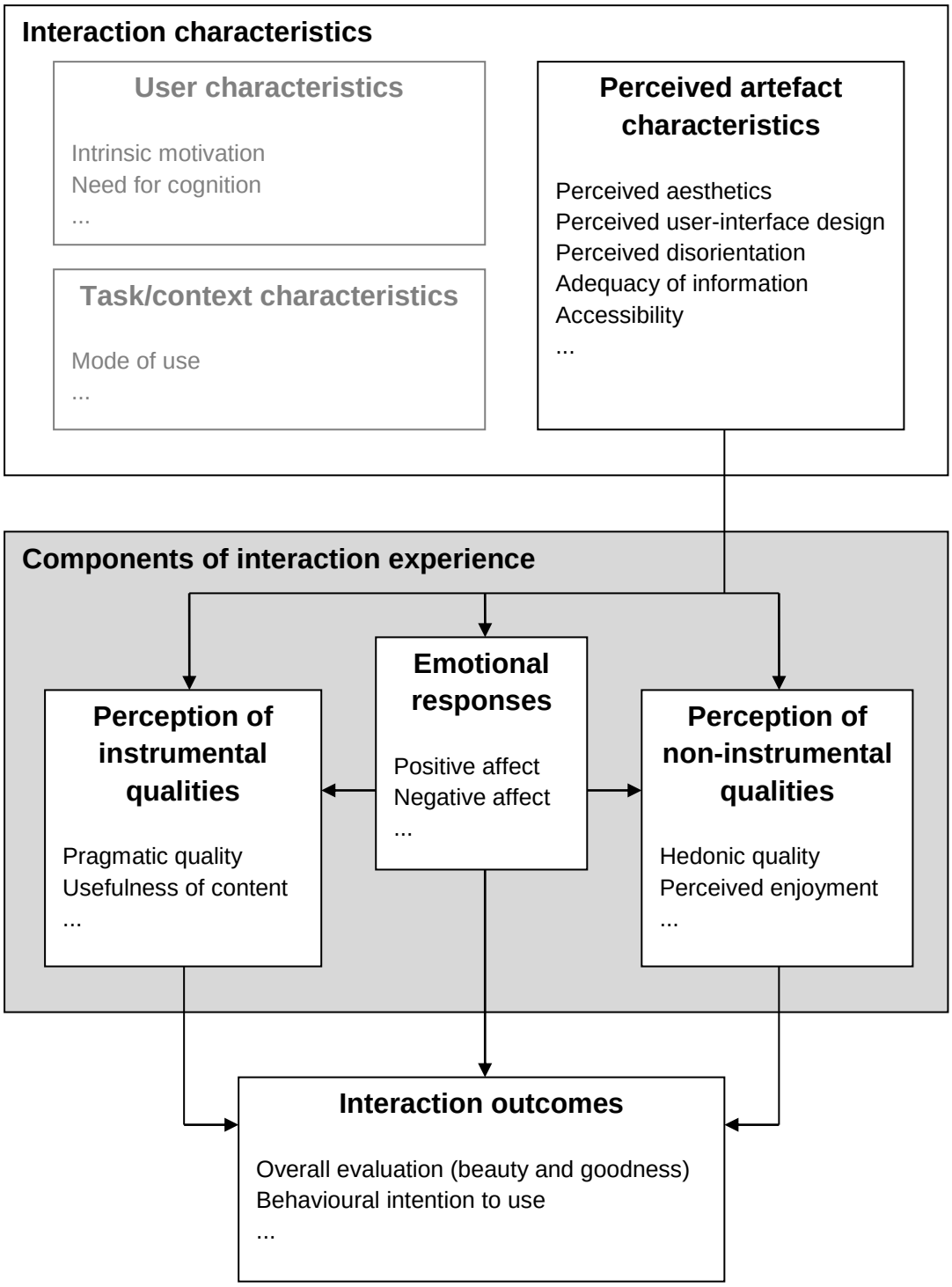


Figure 5.7. An outline of the model of interaction experience with news sites.

First, the paths from artefact-characteristic measures to the affect dimensions were tested. Only two paths were significant: *perceived disorientation* to *negative affect* (β = .32, t = 4.11, p < .001) and *perceived aesthetics* to *positive affect* (β = .32, t = 4.26, p < .001). Therefore, only these paths between artefact characteristics and affective reactions were retained for further analyses.

Next, measures of artefact characteristics were omitted and the paths leading from both affect dimensions to measures of perception of instrumental qualities, perception of non-instrumental qualities, and to *beauty* and *goodness* were tested. The results are presented in Table 5.7. All previously established significant paths of the interaction experience, technology acceptance and integrated experience-acceptance models were also included in the analysis.

Table 5.7. Paths of affect dimensions to experience and outcome measures (artefact characteristics excluded).

Target variable	Predictor variable			
	Positive affect		Negative affect	
	β	t[a]	β	t
Pragmatic quality	.29	***5.63	-.26	***4.60
Hedonic quality	.37	***6.68	-.07	1.35
Usefulness of content	.15	*2.12	-.16	*2.42
Perceived enjoyment	.42	***9.49	-.14	**2.74
Beauty	.16	**3.02	-.04	0.81
Goodness	.09	1.79	-.13	*2.29

[a]Bootstrap, N = 5000.

*p < .05. **p < .01. ***p < .001.

Positive affect was positively and significantly connected to instrumental qualities, non-instrumental qualities, *beauty* judgement and *pragmatic quality*, whereas *negative affect* was more strongly (and negatively) connected to instrumental qualities and *goodness* judgements, and it was also negatively and significantly connected to *perceived enjoyment*. The non-significant paths of *positive affect* to *goodness* and *negative affect* to *hedonic quality* and *beauty* were omitted in further analyses.

Finally, all measures and previously significant paths were re-introduced to the model and a bootstrapping algorithm (N = 5000) was run to eliminate paths suppressed by the presence of other variables. With the inclusion of measures of artefact characteristics, the paths from *pragmatic quality* to *usefulness of content* and *perceived enjoyment*, and the path from *hedonic quality* to *usefulness of content* lost significance. With the inclusion of the path between *hedonic quality* and *perceived enjoyment*, the direct effect of *perceived aesthetics* on *perceived enjoyment* became non-significant. The non-significant paths were omitted, and the PLS algorithm and bootstrapping procedure were re-run to acquire the final parameters of the model. The model parameters of the integrated model of interaction experience with news sites are presented in Table 5.8.

Table 5.8. Model parameters of the model of interaction experience with news sites.

Prediction of variables of interaction-experience components					
Target variable (number of predictors)	Predictor variable	β	t^a	sr^2	Unique effect size (f^2)[b]
Pragmatic quality (4)	Perceived disorientation	-.24	***3.38	.04	.07
R^2 = .43	Perceived U-I design	.44	***6.84	.14	.24
	Negative affect	-.11	*2.11	.01	.02
	Positive affect	.09	*2.00	.01	.01
Usefulness of content (3)	Accessibility	.30	***3.97	.07	.13
R^2 = .49	Adequacy of Information	.47	***6.93	.17	.32
	Negative affect	-.10	*2.07	.01	.02
Perceived enjoyment (4)	Perceived U-I design	.33	***5.70	.08	.17
R^2 = .52	Negative affect	-.13	**3.01	.02	.04
	Positive affect	.38	***9.47	.12	.24
	Hedonic quality	.24	***3.97	.04	.09
Hedonic quality (2)	Perceived aesthetics	.60	***12.15	.30	.51
R^2 = .42	Positive affect	.11	*2.04	.01	.02
Positive affect (1)	Perceived aesthetics	.41	***7.87	R^2 = .16	.20
Negative affect (1)	Perceived disorientation	.29	***4.85	R^2 = .08	.09
Prediction of variables of interaction outcomes					
Target variable (number of predictors)	Predictor variable	β	t^a	sr^2	Unique effect size (f^2)[b]
Goodness (3)	Negative affect	-.12	*2.00	.01	.02
R^2 = .47	Pragmatic quality	.28	***5.05	.06	.11
	Hedonic quality	.49	***9.43	.19	.35
Behavioural intention (2)	Usefulness of content	.32	***4.53	.09	.14
R^2 = .34	Perceived enjoyment	.38	***5.95	.12	.18
Beauty (2)	Positive affect	.15	**3.08	.02	.04
R^2 = .40	Hedonic quality	.56	***10.85	.28	.46

[a]Bootstrap, N = 5000.

[b]f^2: 0.02 - 0.14 small, 0.15 - 0.34 medium and 0.35 - large.

*p < .05. **p < .01. ***p < .001.

Except for the medium effect of *positive affect* on *perceived enjoyment* (f^2 = .24), the two affect dimensions exerted small effect on other endogenous variables in the model (f^2 ranging from .01 to .04); the significant paths were not excluded from the final model on the basis of their low individual contribution to proportions of variance explained in the endogenous variables. Significant paths with low effect size are included to allow for the examination of structural relationships between affect dimensions and other variables. The significant but small effect of *goodness* on *behavioural intention* was removed from the model in order to keep the outcome

measures separate from each other and to predict interaction outcomes only from components of interaction experience. The path diagram of the model of interaction experience with news sites is presented in Figure 5.8.

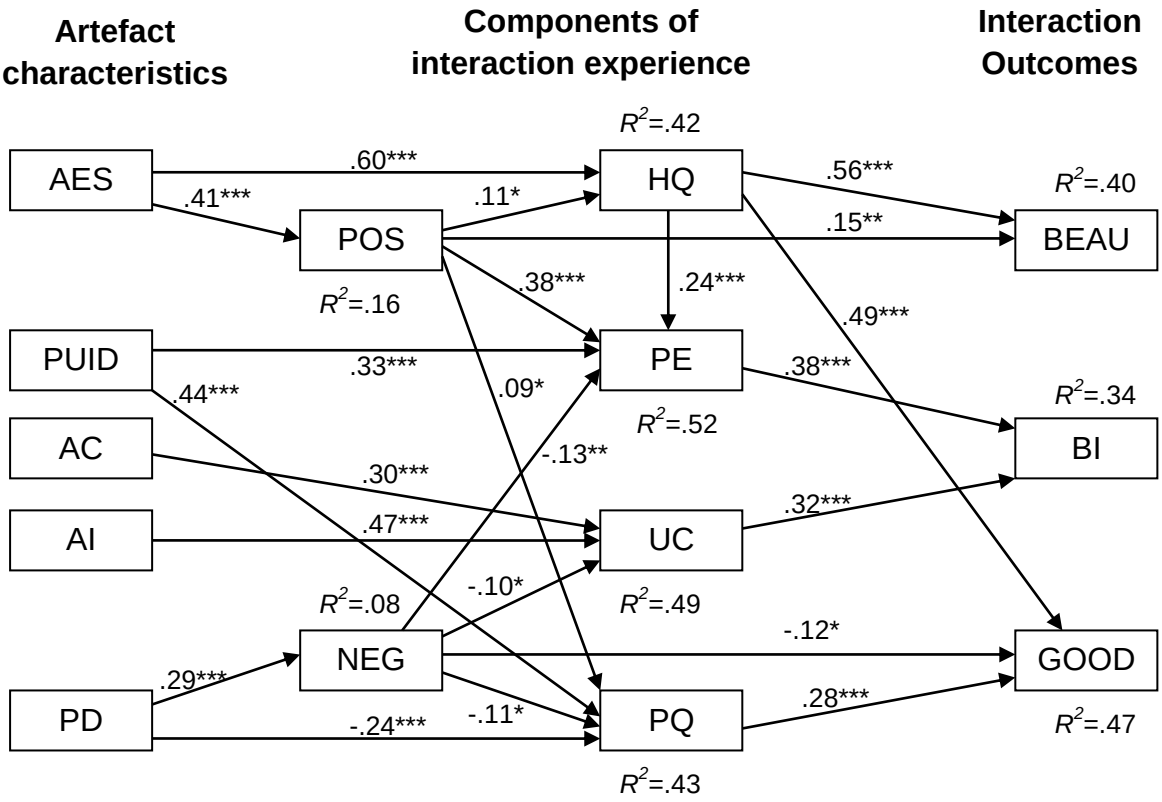


Figure 5.8. Path diagram of the model of interaction experience with news sites.

Note. AES: perceived aesthetics. PUID: perceived user-interface design. AC: accessibility. AI: adequacy of information. PD: perceived disorientation. POS: positive affect. NEG: negative affect. HQ: hedonic quality. PE: perceived enjoyment. UC: usefulness of content. PQ: pragmatic quality. BEAU: beauty. BI: behavioural intention. GOOD: goodness.
*p < .05, **p < .01, ***p < .001

5.3.5. Differences between users of BBC and other news sites

As discussed in Chapter 4, participants of Study 2 were invited to use any news site of their choosing before completing the online interaction-experience questionnaire. Two-thirds of participants (66%) chose to use the BBC site (see Chapter 4, Table 4.2). In order to explore the differences in model parameters between users of BBC and other news sites, the data was split into two: users of the BBC news site (n = 202) and users of any other news sites (n = 103). Note the exploratory nature of the analysis based on this split of the data. The artefact in the group of users of other news sites is comprised of a wide range of sites and these differences are not accounted for. However, because BBC is the most prestigious news provider in the United Kingdom, it is reasonable to examine it separately from and in comparison with the use of other news sites. The separate examination of the BBC site from other news sites is expected to shed light on the applicability of the model to a particular news site. However, because BBC users provided two thirds of the data, the model is biased to a great extent by answers of BBC users. Nevertheless, the sample sizes after the split was adequate for PLS analysis on both groups. The significance of differences between model parameters for the BBC and other news sites were calculated using t tests based on the β coefficients and their associated standard errors acquired by running a bootstrapping procedure (Chin, 2000). The model parameters for the two groups with paths hypothesised on the basis of the joint model and the tests of differences of path coefficients are presented in Table 5.9.

Table 5.9. Differences of standardised path coefficients between BBC and other news sites.

Prediction of variables of interaction-experience components									
Target variable	Predictor variable	BBC *n* = 202			Other sites *n* = 103			BBC - Other *df* = 303	
		β	*t*[a]	*SE*	β	*t*	*SE*	Δβ	|*t*|
PQ	PD	-.35	***4.51	.08	-.07	0.59	.11	-.28	*2.10
R^2_{BBC} = .46	PUID	.34	***4.81	.07	.60	***5.41	.11	-.26	*2.05
R^2_{other} = .44	NEG	-.12	1.80	.07	-.09	0.94	.09	-.03	0.78
	POS	.12	*2.08	.06	.05	0.57	.08	.07	0.70
UC	AC	.33	**3.24	.10	.23	*2.41	.10	.10	0.62
R^2_{BBC} = .54	AI	.47	***5.14	.09	.50	***4.94	.10	-.03	0.23
R^2_{other} = .42	NEG	-.09	1.45	.06	-.09	1.09	.08	.00	0.04
PE	PUID	.27	***3.85	.07	.46	***5.74	.08	-.19	1.74
R^2_{BBC} = .53	NEG	-.09	1.69	.05	-.19	*1.99	.09	.10	1.01
R^2_{other} = .56	POS	.38	***7.97	.05	.34	***4.93	.07	.04	0.45
	HQ	.32	***5.24	.06	.09	0.96	.09	.23	*2.14
HQ	AES	.59	***10.42	.06	.62	***7.48	.08	-.03	0.30
R^2_{BBC} = .41	POS	.13	1.91	.07	.08	0.90	.09	.05	0.42
R^2_{other} = .45									
POS	AES	.31	***4.15	.07	.55	***9.18	.06	-.24	*2.20
R^2_{BBC} = .09									
R^2_{other} = .31									
NEG	PD	.33	***5.04	.07	.25	1.92	.13	.08	0.62
R^2_{BBC} = .11									
R^2_{other} = .06									
Prediction of variables of interaction outcomes									
Target	Predictor	BBC			Other			BBC - Other	
		β	*t*	*SE*	β	*t*	*SE*	Δβ	|*t*|
GOOD	NEG	-.10	1.34	.07	-.18	*2.00	.09	.08	0.48
R^2_{BBC} = .39	PQ	.26	***3.85	.07	.33	***3.76	.09	-.07	0.52
R^2_{other} = .62	HQ	.44	***8.61	.05	.54	***5.65	.10	-.10	0.31
BI	UC	.42	***4.94	.09	.12	0.92	.13	.31	*2.06
R^2_{BBC} = .33	PE	.28	***3.69	.07	.59	***6.10	.10	-.31	*2.51
R^2_{other} = .42									
BEAU	POS	.21	***3.30	.06	.04	0.54	.07	.17	1.69
R^2_{BBC} = .33	HQ	.48	***7.28	.07	.70	***10.08	.07	-.22	*2.15
R^2_{other} = .52									

Note. AC: accessibility. AES: aesthetics. AI: adequacy of information. BEAU: beauty. BI: behavioural intention. GOOD: goodness. HQ: hedonic quality. NEG: negative affect. PD: perceived disorientation. PE: perceived enjoyment. POS: positive affect. PQ: pragmatic quality. PUID: perceived user-interface design. UC: usefulness of content.

[a]Bootstrap, *N* = 5000.

*$p < .05$. **$p < .01$. ***$p < .001$.

The majority of paths remained significant for both groups, except for several paths leading from affect variables to other latent variables. However, these paths represented the smallest effect sizes in the joint model; consequently, these were the most sensitive to the drop in sample size resulting from the data split. *Perceived disorientation* had a significantly higher negative impact on *pragmatic quality* in the BBC group ($\Delta\beta = .28$, $|t| = 2.10$, $p < .05$). In fact, *perceived disorientation* was not a significant predictor of *pragmatic quality* for users of other news sites ($\beta = -.07$, $t = 0.59$, *ns*). However, *perceived user-interface design* had a significantly higher impact on *pragmatic quality* for users of news sites other than BBC ($\Delta\beta = .26$, $|t| = 2.05$, $p < .05$). Furthermore, *hedonic quality* was a significant predictor of *perceived enjoyment* only in the BBC group, and the path coefficients were significantly different in the two groups ($\Delta\beta = .23$, $|t| = 2.14$, $p < .05$).

For the BBC group, additional analysis, motivated by the observed between-group differences in the prediction of outcome measures, showed that the significant path from *goodness* to *behavioural intention*, tested in relation to the integrated experience-acceptance model in Section 5.3.3, became non-significant ($\beta = .08$, $t = 1.17$, *ns*) with the inclusion of a direct path from *pragmatic quality* to *behavioural intention* ($\beta = .17$, $t = 1.96$, $p < .05$). Mediation analysis revealed that the effect of *pragmatic quality* on *behavioural intention* was not mediated through *usefulness of content*. These findings indicate that the product-attribute construct of *pragmatic quality* is a direct predictor of the technology-acceptance outcome measure of *behavioural intention* for users of the BBC site.

There were a number of notable differences between the two groups. *Perceived aesthetics* was a significantly stronger predictor of *positive affect* for news sites other

than BBC ($\Delta\beta$ = .24, $|t|$ = 2.20, p < .05). For other sites, *positive affect* was exclusively related to *perceived enjoyment* and the path from *usefulness of content* to *behavioural intention* lost its significance. *Usefulness of content* had a significantly higher impact on *behavioural intention* in the BBC group ($\Delta\beta$ = .31, $|t|$ = 2.06, p < .05). On the other hand, the impact of *perceived enjoyment* on *behavioural intention* was significantly higher for news sites other than BBC ($\Delta\beta$ = .31, $|t|$ = 2.51, p < .05). The prediction of *behavioural intention* from *perceived enjoyment* alone proved to be better in the sample of user using sites other than BBC than its prediction from *usefulness of content* and *perceived enjoyment* together in the BBC sample. Similarly, Dickinger et al. (2008) found that the effect of *perceived enjoyment* was twice as strong on intention to use as the effect of *perceived usefulness*.

For BBC, *usefulness of content* was a stronger predictor of *behavioural intention* than *perceived enjoyment*. *Hedonic quality* had a significantly higher impact on *beauty* judgements for news sites other than BBC ($\Delta\beta$ = .22, $|t|$ = 2.15, p < .05). Overall, all variables of interaction outcomes were better predicted in the case of news sites other than the BBC site. For the BBC site, pragmatic aspects were stronger predictors of *behavioural intention* than hedonic aspects. For other news sites, hedonic aspects dominated the prediction of all outcome variables.

Independent *t* tests were conducted using the latent variable scores of all collected measures to examine the differences between the responses to the BBC site and other news sites. Users of the BBC site (M = 4.80, SE = 0.08) reported lower levels of *perceived enjoyment* than users of other news sites (M = 5.27, SE = 0.13). This difference was significant (t (303) = -3.15, p < .01) and it represented a small effect size (r = .18). Users of the BBC site (M = 3.66, SE = 0.08) have also reported lower

levels of *positive affect* than users of other news sites (M = 3.95, SE = 0.13). This difference was significant (t (303) = -2.02, $p < .05$) and it represented a small effect size (r = .12). No significant differences were detected on any other variables between the two groups.

5.4. Discussion

To test hypotheses regarding the structural relationships between the variables selected for the current study, three models were specified, based on the relevant literature: an interaction experience model, a technology acceptance model and an integrated experience-acceptance model. All three models were supported and the majority of hypotheses were confirmed (see Tables 5.1, 5.3 and 5.5). A model of interaction experience with news sites was formulated and tested, based on the components of user-experience model, which included all variables from the previous models. Differences of path coefficients in the model of interaction experience with news sites between users of BBC and users of other news sites were tested.

5.4.1. Interaction experience model

In the interaction experience model, *goodness* judgement of an artefact was affected by both *pragmatic quality* and *hedonic quality*, whereas *beauty* was not affected by *pragmatic quality*. Notably, the effect size of *hedonic quality* on *goodness* was more than twice than that of *pragmatic quality* (f^2 = .33 and .14, respectively), suggesting that *goodness* judgements of news sites are predominantly grounded in hedonic aspects. Product attributes, in turn, were predicted from measures of perceptions of artefact characteristics. *Perceived disorientation* was negatively, and *perceived user-interface design* was positively related to *pragmatic quality*, together accounting for 42% of its variance. *Perceived aesthetics* alone accounted for 41% of variance in

hedonic quality, but it was also a significant direct predictor of *beauty*, with its effect partially mediated through *hedonic quality*. In the model, 46% variance in *goodness* and 42% variance in *beauty* was accounted for by predictors. Although *positive affect* was positively and significantly related to *hedonic quality*, and *negative affect* was negatively and significantly related to *pragmatic quality*, affect variables were not included in the final model, because of their small effect sizes.

The amount of variance explained in endogenous variables was markedly lower than in the think-aloud study (see Chapter 3, Section 3.6.2). In the regression analyses of the think-aloud study, 84% of variance in *goodness* was accounted for by *pragmatic quality* and *hedonic quality*, and 50% of variance in *beauty* was accounted for by *hedonic quality* alone. However, it is important to note that the two studies were different in several respects. Firstly, the think-aloud study took place in a laboratory setting, whereas in the current study participants used a news site at a location of their own selection in an uncontrolled setting, without the presence of an experimenter. Secondly, the think-aloud study used one particular news site, whereas the current study allowed participants to use a news site of their choice. Thirdly, 15 of the 25 participants of the think-aloud study were not familiar with the particular news site that they were asked to use, while in the current study, participants were already users of the news site of their selection. Generally, the think-aloud study was more controlled with respect to its design and procedure. Nevertheless, the hypotheses for the interaction experience model were supported, indicating the robustness of the model.

Similarly, *classical aesthetics* and *expressive aesthetics* accounted for 71% of variance in *hedonic quality* in the think-aloud study, but 42% of variance was

accounted for by *perceived aesthetics* in *hedonic quality* in the current study. In the think-aloud study, 67% of variance in *pragmatic quality* was accounted for by *classical aesthetics*. In the current study, a composite of the two aesthetics dimensions was used, based on psychometric considerations. Instead of *classical aesthetics*, *perceived user-interface design* was a strong and significant predictor of *hedonic quality*, while the effect of *perceived aesthetics* on *pragmatic quality* was non-significant. However, when *perceived user-interface design* was removed from the model, the effect of *perceived aesthetics* on *pragmatic quality* became significant. These results suggest that *perceived aesthetics* and *perceived user-interface design* share a significant portion of variance, presumably the variance of *classical-aesthetics* items in the composite *perceived aesthetics* measure. Furthermore, the measures of *perceived aesthetics* and *perceived user-interface design* were strongly correlated ($r = .60$, $p < .01$). In the technology acceptance model, both *perceived aesthetics* and *perceived user-interface design* had a significant direct effect on *perceived enjoyment*. However, in the model of interaction experience with news sites, *perceived aesthetics* lost its direct effect on *perceived enjoyment*; its effect was mediated through *hedonic quality* and *positive affect* (see Figure 5.8). A high amount of shared variance was shared between *perceived aesthetics* and *perceived user-interface design*. *Perceived user-interface design*, which was selected to address the measurement of the *layout* category of the think-aloud study, is therefore regarded to be conceptually similar to *classical aesthetics*, and as such it is positively related to *pragmatic quality* (i.e., user-perceived usability).

5.4.2. Technology acceptance model

In the technology acceptance model, 39% of variance in *behavioural intention* was accounted for by *perceived enjoyment*, *usefulness of content* and *perceived user-*

interface design. This result is consistent with the findings of four longitudinal studies of Venkatesh and Davis (2000) applying the extended version of the technology acceptance model (TAM2; see Chapter 2), where 34%-52% of variance was explained by technology acceptance constructs. Together, *adequacy of information*, as a measure of information quality, and *accessibility*, as a measure of service quality, accounted for 48% of the variance in *usefulness of content*, which was used as a proxy for *perceived usefulness* in the current study. In the four longitudinal studies of Venkatesh and Davis (2000), the variance explained in perceived usefulness in voluntary use setting ranged from 40% to 60%. Although they used different variables to predict perceived usefulness, the variance explained in the current study in *usefulness of content* falls between these values. *Perceived enjoyment*, the other technology acceptance construct in the model, was predicted from *perceived user-interface design*, *perceived aesthetics*, *positive affect* and *negative affect*, with the latest being negatively related to it. Together these variables accounted for 52% of variance in *perceived enjoyment*. By comparison, in the study of Cyr and colleagues (2006), 43% of variance in *perceived enjoyment* was accounted for by *design aesthetics* and *perceived ease of use*. *Perceived user-interface design*, however, was also directly related to *behavioural intention*, with a slightly bigger unique effect size (f^2 = .08) than *usefulness of content* and *perceived enjoyment* (f^2 = .07 for both). This finding underlines the importance of the classically aesthetic design and layout of the interface in the acceptance of news sites, because all the functionalities of Web sites are presented via their user-interface. *Perceived user-interface design* significantly contributed to accounting for variance in *perceived enjoyment*, but its variance shared with aspects of user-perceived usability, as demonstrated in the analysis of the interaction experience model, remained

unexplored in the current technology acceptance model, because *perceived ease of use* was not measured in this study. It may be that *perceived ease of use* (together with *perceived enjoyment*) would fully mediate the direct effect of *perceived user-interface design* on *behavioural intention*, consistent with the findings of Cho and colleagues (2009; see Chapter 2, Figure 2.5.).

5.4.3. Integrated experience-acceptance model

An integrated experience-acceptance model was specified to examine the relationships between technology acceptance constructs and product attributes, with measures of perceived product characteristics and affect excluded from the analysis. Product attributes were significant predictors of technology acceptance constructs, and *pragmatic quality* and *hedonic quality* accounted for 32% of variance in *perceived enjoyment* and 19% of variance in *usefulness of content*. *Pragmatic quality* in the model had a significant effect on *perceived enjoyment*, whereas in the study of van Schaik and Ling (2011), *pragmatic quality* was a determinant of *perceived ease of use*. Because *perceived ease of use* is regarded as an antecedent of *perceived enjoyment* (Cyr et al, 2006, 2007), and no measure of *perceived ease of use* was included in the current study, it is possible that the direct effect of *pragmatic quality* on *perceived enjoyment* is due to the lack of the mediating effect of *perceived ease of use*. Although van Schaik and Ling (2011) found that product evaluations are not direct determinants of *intention to use*, the current study found a significant positive link between *goodness* and *behavioural intention*; however, its effect size was small (f^2 = .03). With the inclusion of measures of artefact characteristics and affect, only the paths from *hedonic quality* to *usefulness of content*, and from *pragmatic quality* to *perceived enjoyment* and *usefulness of content* became non-significant. *Behavioural intention* was a significant predictor of

frequency of use, but it only accounted for 1% of unique variance when *baseline use-frequency* was included in the model. Participants were already users of the sites they chose to use before completing the questionnaires, and the analysis showed that their frequency of use of a particular site was predominantly related to their frequency of use of news sites in general. This finding suggests that frequency of use of a particular news site is not a useful measure for actual use behaviour. For example, preference to use a particular news site over other news sites or other sources of information may be more useful in connecting *behavioural intention* to actual behaviour in the domain of news site use. Nevertheless, the correlation between *behavioural intention* and *frequency of use* was medium-sized ($r = .33$, $p < .01$), and *behavioural intention* alone accounted for 11% of variance in *frequency of use*.

5.4.4. A model of interaction experience with news sites

Following the tests of hypotheses, a model of interaction experience with news sites was constructed, based on the components of user-experience model by Thüring and Mahlke (2007). Accordingly, the measures were arranged in three groups: interaction characteristics, components of interaction experience and interaction outcomes. Interaction characteristics included measures of perceived artefact characteristics, namely: *accessibility*, *adequacy of information*, *perceived disorientation*, *perceived user-interface design* and *perceived aesthetics*. These measures were used to predict variables of components of interaction experience, which were arranged in three groups: (a) perception of instrumental qualities, including *pragmatic quality* and *usefulness of content*, (b) perception of non-instrumental qualities, including *hedonic quality* and *perceived enjoyment*, and (c) emotional responses (which only included measures of affective reactions in the

current study), including *positive affect* and *negative affect*. The overall quality judgements of *beauty* and *goodness*, and *behavioural intention* were grouped together as measures of interaction outcomes (see Figure 5.7).

Tests of hypotheses in the interaction-experience and technology-acceptance models provided evidence for the connections between perceived artefact characteristics and perceived instrumental and non-instrumental qualities. Perceived artefact characteristics were used to predict emotional responses. The analysis revealed that *perceived aesthetics* is a significant predictor of *positive affect* and it accounted for 16% of its variance. *Perceived disorientation* was significantly and poitively related to *negative affect*, accounting for 8% of its variance. Emotional responses, in turn, were used to predict the variables of perceptions of instrumental and non-instrumental qualities. Analyses revealed that *negative affect* was significantly and negatively related to *perceived enjoyment*, *usefulness of content*, *pragmatic quality* and *goodness*. *Positive affect*, on the other hand, was found to be positively and significantly related to *hedonic quality*, *perceived enjoyment*, *pragmatic quality* and *beauty*. In short, positive affect was predominantly positively connected to the perception of non-instrumental qualities and beauty evaluation, and negative affect was predominantly negatively connected to the perception of instrumental qualities and goodness evaluation. These findings lend supports to the distinction between hygiene- and motivator factors as determinants of interaction experience (Zhang & von Dran, 2000). *Negative affect* experienced during the interaction negatively affected perceptions of instrumental qualities and goodness evaluations, reducing the quality of experience, whereas *positive affect* was positively related to non-instrumental qualities and beauty evaluations, promoting the quality of experience. Furthermore, the connection of affect dimensions to perceived product

characteristics helps to clarify sources of positive and negative affect that are rooted in system design. For example, better information architecture is expected to result in lower levels of disorientation perceived by users of a system, which will consequently lead to lower levels of *negative affect* and higher levels of perceptions of instrumental qualities and *goodness* judgements. Increasing the quality of aesthetic design of a system or interface is expected to result in higher levels of *positive affect* during interaction, increasing perceptions of non-instrumental qualities and leading to higher values of *beauty* judgements.

5.4.5. Implications

Potential design guidance based on the model for a particular system needs to take into account the magnitude of effect of a particular perceived artefact characteristic on a target variable of interest, as well as the extent to which the perceived artefact characteristic in question is rated sufficiently high by users or potential users of the system. In other words, the impact of each predictor variable needs to be mapped against the system's performance on the particular variable (Martensen & Grønholdt, 2003). For example, *perceived aesthetics* exerts a strong impact on *hedonic quality* in the model (high path coefficient and large effect size). If a particular news site scores relatively low on *perceived aesthetics*, the recommendation can be made to increase the aesthetic quality of the site. However, if the aesthetic quality of the site is already sufficiently high, then such recommendation is futile, but designers can be advised to maintain quality on that dimension. Conversely, relatively unimportant predictors of measures of interaction experience may score low for a particular site, thus a recommendation can be made to improve the related aspect of the system, but with the condition that changes are not expected to impact the quality of experience substantially.

With regards to emotions in the model, it is important to note that measuring affect is only one way to approach the examination of emotional responses during interaction. Measuring multiple components of emotions (e.g., cognitive appraisals and physiological activation) is expected to result in a more detailed description of the role of emotional responses in interaction experience (Mahlke and Minge, 2008), but it may come at the expense of parsimony and complicating measurement in practical terms. The current model does not aim to provide a detailed description of the role of emotions in interaction experience; it merely considers affective reactions as a component of interaction experience and describes its relationships with other variables.

In the current model, variables of interaction characteristics only include measures of perceived artefact characteristics. However, user characteristics and task/context characteristics are considered to be important aspects of interaction experience (Finneran & Zhang, 2003) and their role needs to be addressed in further research. In Figure 5.7, proposed measures of interaction characteristics are presented in grey; their presence is acknowledged, but their measurement was not included in the current study. Several measures of person characteristics can be considered for inclusion if they were found to be connected to use of Web sites in previous studies. Examples are *spatial ability* (Juvina & van Oostendorp, 2006), *need for cognition* (Amichai-Hamburger et al., 2007) and *intrinsic motivation* (van Schaik & Ling, [a]). Demographic properties of participants, such as age and level of education, may also be included among person characteristics to form different groups of users, for example, to specify ‘target audiences’.

A task/context characteristic that may be involved in further studies is *mode of use*. According to Hassenzahl (2003), users of interactive systems in goal mode are pursuing specific goals during their use and their focus is on the attainment of a desired outcome, whereas in action mode their focus is on the action itself of using the system. Research suggests that users' perceptions of a particular system, and consequently their experience, is influenced by the mode in which they use the system (Hassenzahl & Ullrich, 2007; van Schaik & Ling, 2009, 2011). In the case of news-site use, participants in goal mode may look for specific pieces of information and search for certain types of content. Good organisation of content and links with high information scent are expected to aid them in finding desired content. In action mode, users may browse a news site to see if anything catches their attention. In this case, exploration of the site may be promoted, for example, by an aesthetically stimulating graphical environment and interesting headlines. Everyday use of news sites is expected to be in mixed mode, that is, mode of use is likely to alternate between action and goal mode. For example, exploration behaviour may trigger a search for specific pieces of information, and once a specific interest of a particular user is fulfilled, he/she may revert to browsing for potentially interesting articles.

In summary, the model of interaction experience with news sites incorporates constructs of perceived instrumental qualities, non-instrumental qualities and affective reactions to predict overall evaluations of news sites by users and intention to use. The model aims to account for technology acceptance and overall quality judgements at the same time, including affective reactions in an attempt to tap emotional aspects of interaction experience. Although affective reactions were significantly related to other constructs of interaction experience and outcome measures, their direct contribution to prediction was relatively small. However, there

was a stronger indirect effect of *positive affect* on both overall judgement and technology acceptance. Furthermore, in system design promoting positive affective reactions and preventing negative reactions in users may be a goal in its own right. The relative independence of outcome variables from each other, in terms of being predicted from a different set of variables, supports the need to apply various outcome measures and a sufficiently wide range of components of interaction characteristics to predict them. The model also aims to link components of interaction experience to interaction characteristics, and in particular, to perceptions of artefact characteristics that can be directly connected to and controlled by system design. However, without further empirical testing, the model's power to draw cause-effect relationships between constructs remains limited. To address the need for further support, the next chapter presents an experimental test of some of the aspects of the model.

Chapter 6

An experimental study of interaction experience with news sites

Contents

Chapter 6: experimental study

6.1. Overview

This chapter presents an experimental study of interaction experience with news sites. The experiment was conducted in a laboratory setting at Teesside University with two conditions. In one condition, participants used an adopted news site, in other words, a site that they had already been using regularly before their participation in the experiment. Participants in the other condition used a non-adopted news site, in other words, a site that they had never used before. The main aim of the experiment was to test the model of interaction experience with news sites, specified in the previous chapter, in a controlled setting. The experimental conditions were included to test the model's stability across different stages of adoption. The chapter begins with the presentation of the design of the experiment and the measurement model. Tests of differences in experience between the two experimental conditions are then presented. Following the tests of hypotheses developed in the previous chapter, general-to-specific inference-perspective models are constructed to test the relationship *beauty* and perceived usability, and to test the independence of *pragmatic quality* and *hedonic quality*. The model is extended with *trust* as another aspect perceived instrumental quality and *satisfaction* as an additional interaction outcome, and parameters of the extended model are calculated. Differences in model parameters between the experimental conditions are presented and discussed. Finally, the person-characteristic measures of *centrality of visual product aesthetics* and *spatial ability* are introduced and tested in the model.

6.2. Method

6.2.1. Design

An independent measures design was used with two conditions in a laboratory setting. The independent variable was the news site used during the study. Participants of one group used the BBC news Web site, whereas the other group used The New Zealand Herald (NZH) during the experiment. Outcome measures included variables of interaction experience used in Study 2, complemented with additional variables of perceptions of instrumental qualities and interaction outcomes. Furthermore, person-characteristic variables were included to test for their effects on the level of interaction-experience variables.

6.2.2. Participants

Adult, English-speaking people were recruited for the experiment, who used the BBC news site on a regular basis. Every participant in The NZH condition confirmed that they never used the New Zealand Herald before. In order to support the quality of collected data, recruitment primarily aimed at staff, research students and master's students at Teesside University. Participants received £8 cash compensation for their time and effort. E-mails were circulated by each department of the university among staff and by the university's Graduate Research School among graduate students. A recruitment Web page was launched, where those interested in participation could sign up to one of the ten time slots available[19]. Descriptive statistics of the sample are presented in Table 6.1.

[19] A maximum of 20 participants could sign up for each time-slot.

Table 6.1. Descriptive statistics of demographics, Internet use and news-site use.

Age	Mean	Minimum	Maximum	*SD*	*n*/*N*
BBC	33.07	19	72	10.00	43
NZH	33.76	19	57	10.03	42
Total	33.41	19	72	9.96	85

Gender	Male Percentage (*n*)	Female Percentage (*n*)
BBC	42% (18)	58% (25)
NZH	43% (18)	57% (24)
Total	42% (36)	58% (49)

Highest Level of education	BBC Percentage (n)	NZH Percentage (n)	Total Percentage (n)
GCSE or equivalent	2.3% (1)		1.2% (1)
NVQ Level 1 or 2		4.8% (2)	2.4% (2)
Two or more A-levels	2.3% (1)	4.8% (2)	3.5% (3)
NVQ Level 3, 4 or 5	4.7% (2)	7.1% (3)	5.9% (5)
Degree	27.9% (12)	26.2% (11)	27.1% (23)
Higher degree	62.8% (27)	57.1% (24)	60.0% (51)
News-site use frequency			
Weekly	7.0% (3)	2.4% (1)	4.7% (4)
2-3 times per week	20.9% (9)	21.4% (9)	21.2% (18)
4-6 times per week	16.3% (7)	16.7% (7)	16.5% (14)
Daily	32.6% (14)	26.2% (11)	29.4% (25)
2-3 times per day	16.3% (7)	21.4% (9)	18.8% (16)
More than 3 times per day	7.0% (3)	11.9% (5)	9.4% (8)

Internet use (years)			
Mean	13.00	12.45	12.73
SD	3.56	3.16	3.36
Internet use per week (hours)			
Mean	28.42	27.76	28.09
SD	14.52	16.96	15.68
News-site use duration per day (minutes)			
Mean	44.47	42.10	43.29
SD	79.40	51.23	66.59
News-site use duration per visit (minutes)			
Mean	28.47	17.57	23.08
SD	50.34	10.95	36.82

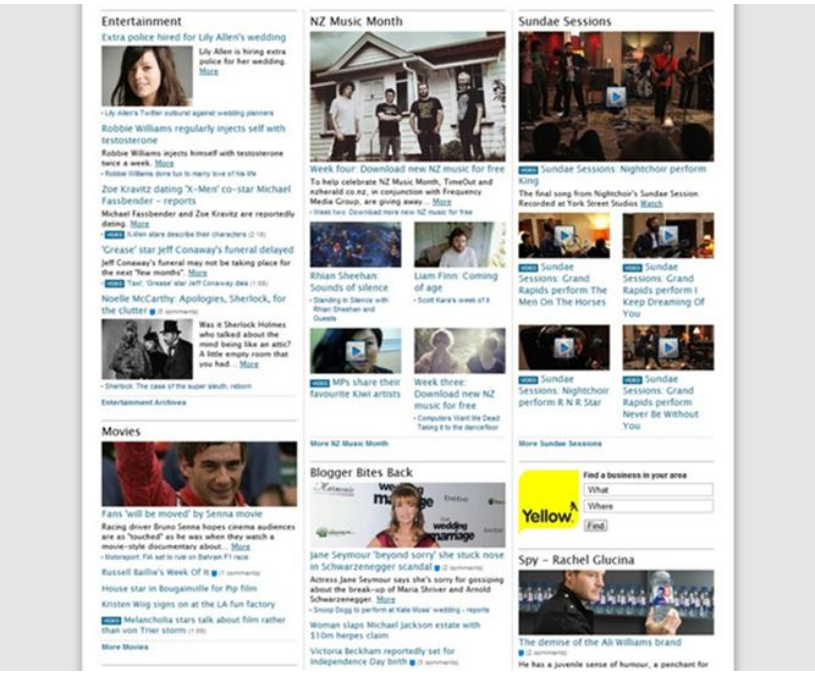
Eighty-five people participated in the experiment (49 females and 36 males), with a mean age of 33.41 years (*SD* = 9.96). Mean experience of using the Internet was 12.73 years (*SD* = 3.36). Mean time of Internet use per week was 28.09 hours (*SD* = 15.68). Mean time of using news sites per day was 43.29 minutes (*SD* = 66.59), and mean time of using news sites per visit was 23.08 minutes (*SD* = 36.82). Values of standard deviations for news-site use duration questions exceeded the means, which

indicates a great variability among participants. However, 50 participants (59%) reported to use news sites for the duration between 10 and 20 minutes per visit. Forty-nine participants (58%) reported to use news sites daily or more frequently. Although participants of the two groups were from the same population, demographic statistics are presented separately for the two groups in Table 6.1; consequently, it is possible to provide evidence for random assignment to groups. Random assignment was successful in that no significant between-group differences were found in demographic, Internet-use and news-site use details.

6.2.3. Materials and equipment

Bespoke experimental software was developed and used for presenting the experiment and collecting responses, coded in Visual Studio .NET 2008. The experimental software was run on identical HP Compaq 8000 Elite SFF personal computers (CPU: Intel Core2Duo E7500, 2.93 GHz; Memory: 3072 RAM; OS: MS Windows 7 Enterprise 32-bit; Monitor: 17'' TFT, 1280x1024 resolution, 75Hz refresh rate). The BBC news Web site (*http://www.bbc.co.uk*) was selected as the adopted news site for the study, because BBC is one of the most prestigious news portals in the United Kingdom. Moreover, two thirds of participants in Study 2 choose to use BBC when asked to use a news site of their choice. The New Zealand Herald (*http://www.nzherald.co.nz*) was selected for the study as a non-adopted site, because it is an English-language news site that participants in the United Kingdom were not likely use, and it is among the most-visited news sites in New Zealand. Examples of screen-shots of the BBC and NZH news sites are presented in Figure 6.1.

a. The BBC homepage (top of page)



b. The NZH entertainment page (scrolled down)

Figure 6.1. Screen-shots of the news sites used in Study 3, as of 3 June 2011.

Because the primary aim of the experimental study was to test the model of interaction experience with news sites (see Chapter 5, Section 5.3.4), measures of

interaction outcomes and components of interaction experience from Study 2 were also used in the current study. The 10-item abridged version of the AttrakDiff2 questionnaire (Hassenzahl et al., 2010) was used to measure perceptions of *pragmatic quality* and *hedonic quality* (four items each), and *beauty* and *goodness* evaluations of the sites (one item each). A 3-item questionnaire adopted from Sun and Zhang (2008) was used to measure *perceived enjoyment*. *Usefulness of conte*nt was measured with a 3-item scale adopted from Yang and colleagues (2005). *Behavioural intention* was measured by a 2-item scale with separate wording for the two experimental conditions. Affective reactions were measured using the 20-item positive and negative affect schedule (Watson et al., 1988). As measures of perceived artefact characteristics, the 7-item *perceived disorientation* scale (Ahuja & Webster, 2001), the 3-item measure of *adequacy of information* (Yang et al., 2005) and the 8-item *perceived aesthetics* scale (Porat & Tractinsky, in press) were used in the study.

In addition to measures applied in Study 2, four new variables were included[20]. A 4-item measure of *trust* was added to measures of perceptions of instrumental qualities (*usefulness of content* and *pragmatic quality*). The items were worded based on the dimensions of trust in news media (Kohring & Matthes, 2007). A 3-item measure of *satisfaction* was added to the indicators of interaction outcomes, adopted from (Martensen & Grønholdt, 2003). *Centrality of visual product aesthetics* (*CVPA*) and *spatial ability* were included in the study as person-characteristic measures. *CVPA* was measured with an 11-item scale adopted from Bloch and colleagues (2003).

[20] The rationale for including the new variables is presented in Sections 6.4.4.1, 6.4.4.2, 6.4.7.1 and 6.4.7.2.

Spatial ability of participants was assessed with ten trial problems of the mental-rotation test developed by Vandenberg and Kuse (1978). Items of each scale used in the study are presented in Appendix 6.1.

6.2.4. Procedure

Upon their arrival in a computer lab at Teesside University, participants were randomly assigned to one of the experimental conditions and received general instructions verbally from the researcher. Following this, they read the participant's information sheet and completed the consent form (see Appendix 6.2). All further instructions and material were presented on the computer screen. The person-characteristic measures were administered before the participants were presented the news site. The mental-rotation test and *CVPA* were followed by a self-paced exposure of five screen shots of the site (BBC or NZH, depending on the experimental condition). The screen shots were taken on pages with similar content from the news sites (home page at top, news page at top, news page scrolled down, sports page scrolled down and entertainment page scrolled down) and were presented in the same order for the two conditions. The screen shots were taken on different pages to provide a cross-section of the sites rather than just a picture of the home page. Following the screen-shots, participants filled in AttrakDiff2 (T1 – after presentation). Attrakdiff2 was administered again after the participants used the site for ten minutes (T2 – after interaction). All remaining measures were also administered after interaction with the site. (See Appendix 6.1 for the presentation order of the measures.) Participants worked individually and they could not leave the experimental software until they had provided a full response to every question. The full procedure took approximately 30 minutes to complete.

6.3. Measurement model

6.3.1. Perceived aesthetics

In Study 2, the factor structure of the two-dimensional *perceived aesthetics* scale did not reproduce statistically; therefore, a unidimensional scale was applied (see Chapter 4, Section 4.4.2.1). In order to decide whether *perceived aesthetics* should be treated as a unidimensional or a two-dimensional scale in Study 3, PLS analysis was conducted using factor-weighting scheme with *classical aesthetics*, *expressive aesthetics* and a combined *perceived aesthetics* scale as predictor variables, and case number as outcome variable[21]. The item loadings and cross-loadings of perceived aesthetics items, average variance extracted (*AVE*) in the aesthetics scales and composite reliability (*CR*) values of the scales are presented in Table 6.2.

Table 6.2. Loadings and cross-loadings of perceived aesthetics items, and values of average variance extracted in aesthetics scales and composite reliability.

Latent variable	Item	Classical aesthetics *AVE* = .62 *CR* = .86	Expressive aesthetics *AVE* = .75 *CR* = .92	Perceived aesthetics *AVE* = .58 *CR* = .91
Classical aesthetics	Clean	0.89	0.47	0.73
	Pleasant	0.92	0.65	0.86
	Symmetrical	0.51	0.33	0.46
	Aesthetic	0.75	0.66	0.77
Expressive aesthetics	Original	0.52	0.88	0.77
	Sophisticated	0.60	0.86	0.81
	Spectacular	0.58	0.85	0.79
	Creative	0.62	0.87	0.82

The square of the correlation between *classical aesthetics* and *expressive aesthetics* (r^2 = .44) was lower than their *AVE* scores (.62 and .75, respectively), supporting

[21] Case number was included because the PLS algorithm can only be calculated on well-formed path models. Case number was used as the outcome variable, because it was expected (and proved) to be independent of other variables.

discriminant validity. However, the two scales were strongly and significantly correlated (r = .67, p < .001), and item cross-loadings were high, ranging between .33 and .66. Values of composite reliability, as a measure of internal consistency, were satisfactory for both the two-dimensional and unidimensional solutions. The unidimensional solution explained 58% of variance in all eight items and exhibited high internal consistency. The only item with lower than .70 loading was *symmetrical*, which was also the item with the lowest loading and cross-loading in the two-factor solution. Because of high cross-loadings in the two-dimensional solution (low item-level reliability) and in order to facilitate comparability with Study 2, a unidimensional *perceived aesthetics* scale was used in Study 3.

6.3.2. Reliability and convergent validity

A measurement model was tested by assigning measurement items to latent variables and drawing all possible links between latent variables, with the inner-weighting option set to factorial scheme in SmartPLS (Chin, 2010). Coefficients of discriminant validity and reliability, and item loadings are presented in Table 6.3. Cross-loadings are presented in Appendix 6.3.

In the *perceived aesthetics* scale, item CA3 (symmetrical) had a loading lower than .50. However, this loading was significant, and the removal of the item left the composite reliability value of the scale unaffected; therefore, CA3 was retained in the measurement model. The loadings of items CVPA07 and CVPA08 on the *CVPA acumen* scale were low and non-significant. As a consequence of removing these two items from the measurement model, composite reliability of the *CVPA acumen* scale increased to .88 from .80, and average variance extracted for the scale increased from .51 to .79.

Table 6.3. Coefficients of reliability and convergent validity.

Construct/indicator	Average variance extracted	Composite reliability	Loading	Standard error	t^a
Perceived aesthetics	0.58	0.92			
- CA1			*0.69	0.10	7.13
- CA2			0.83	0.03	27.50
- CA3			*0.47	0.12	4.04
- CA4			0.79	0.05	15.57
- EA1			0.78	0.06	13.83
- EA2			0.81	0.05	16.57
- EA3			0.82	0.04	21.24
- EA4			0.84	0.04	22.53
Perceived disorientation	0.57	0.90			
- PD1			0.81	0.06	14.08
- PD2			0.73	0.09	8.49
- PD3			*0.67	0.09	7.25
- PD4			0.71	0.08	9.12
- PD5			0.82	0.04	20.38
- PD6			0.82	0.05	15.48
- PD7			0.72	0.06	12.98
Adequacy of information	0.81	0.93			
- AI1			0.90	0.03	34.72
- AI2			0.90	0.02	38.41
- AI3			0.89	0.03	26.02
Usefulness of content	0.62	0.83			
- UC1			0.72	0.11	6.68
- UC2			0.87	0.05	16.23
- UC3			0.78	0.04	18.53
Trust	0.75	0.92			
- TRU1			0.80	0.08	9.52
- TRU2			0.91	0.05	19.57
- TRU3			0.91	0.03	35.88
- TRU4			0.84	0.05	18.55
Perceived enjoyment	0.86	0.95			
- PE1			0.95	0.01	84.07
- PE2			0.96	0.01	110.33
- PE3			0.88	0.05	16.19
Behavioural intention	0.98	0.99			
- BI1			0.99	0.00	239.93
- BI2			0.99	0.00	209.25
Satisfaction	0.84	0.94			
- SAT1			0.94	0.01	65.25
- SAT2			0.90	0.03	32.47
- SAT3			0.91	0.02	45.97
CVPA Value	0.70	0.90			
- CVPA01			0.79	0.07	10.76
- CVPA02			0.86	0.06	15.47
- CVPA03			0.90	0.03	33.08
- CVPA04			0.79	0.06	14.21
CVPA Acumen	0.51	0.80			
- CVPA05			0.88	0.36	2.43

- CVPA06			0.86	0.28	3.05
- CVPA07			*0.54	0.34	1.61
- CVPA08			*0.51	0.32	1.62
CVPA Response	0.71	0.88			
- CVPA09			0.72	0.12	6.07
- CVPA10			0.88	0.05	18.70
- CVPA11			0.91	0.04	24.35
Pragmatic quality T1	0.53	0.80			
- PQ1_1			0.82	0.04	18.33
- PQ2_1			*0.23	0.22	1.07
- PQ3_1			0.86	0.06	15.01
- PQ4_1			0.81	0.05	17.81
Pragmatic quality T2	0.58	0.84			
- PQ1_2			0.88	0.03	30.67
- PQ2_2			*0.50	0.15	3.29
- PQ3_2			0.82	0.09	9.28
- PQ4_2			0.79	0.06	14.19
Hedonic qualityT1	0.63	0.87			
- HQ1_1			0.80	0.05	16.90
- HQ2_1			0.77	0.06	12.71
- HQ3_1			0.74	0.10	7.35
- HQ4_1			0.85	0.03	27.28
Hedonic qualityT2	0.69	0.90			
- HQ1_2			0.84	0.04	19.24
- HQ2_2			0.86	0.03	28.64
- HQ3_2			0.71	0.11	6.50
- HQ4_2			0.91	0.02	41.73
Positive affect	0.55	0.92			
- POS1			*0.69	0.05	13.33
- POS2			0.86	0.03	33.55
- POS3			0.74	0.05	13.53
- POS4			0.78	0.07	10.86
- POS5			0.78	0.05	14.83
- POS6			*0.57	0.09	6.52
- POS7			0.79	0.05	16.31
- POS8			0.78	0.06	13.14
- POS9			*0.63	0.09	7.07
- POS10			0.75	0.06	13.33
Negative affect	0.38	0.85			
- NEG1			*0.50	0.15	3.26
- NEG2			*0.52	0.18	2.95
- NEG3			*0.47	0.19	2.50
- NEG4			*0.69	0.19	3.66
- NEG5			*0.63	0.18	3.59
- NEG6			0.73	0.20	3.65
- NEG7			0.73	0.22	3.39
- NEG8			*0.69	0.18	3.76
- NEG9			*0.42	0.24	1.77
- NEG10			*0.65	0.21	3.06

[a]Bootstrap, *N* = 5000.

*Loading < .70.

Item PQ2 of *pragmatic quality* (unpredictable-predictable) had a low loading in Study 2 (see Chapter 4, Section 4.4.2.2), but this loading was significant and the item was retained in the measurement model. In the current study, the same item proved to have a low loading on *pragmatic quality*, both before and after news-site use (T1 and T2, respectively). Furthermore, the before-use loading of the item was non-significant and very low. Therefore, item PQ2_1 was removed from *pragmatic quality T1*. PQ2_2 was also removed from *pragmatic quality T2* in order to facilitate comparability of *pragmatic quality* scores between T1 and T2 within the study. As a consequence of removing the item, composite reliability increased to .87 for both T1 and T2, and values of average variance extracted increased markedly (T1: from .53 to .69; T2: from .58 to .70).

On average, *negative affect* only retained 38% of variance from its indicators. According to Henseler et al. (2009), average variance extracted should be greater than 50% (*AVE* > .50). Item NEG9 (jittery) had a low and non-significant loading. The removal of NEG9 increased average variance extracted by .02 to .40. Generally, all loadings should be above .70 in a measurement model, which indicates that almost 50% of variance is shared by a particular indicator and the latent variable it loads on. Eight items out of ten in *negative affect* had a loading lower than .70, although two of them (NEG4 and NEG8) were only slightly below this value. By comparison, *positive affect* had seven items loading above .70, and apart from the previously removed items, the rest of the measurement model had only two loadings slightly lower than .70. Furthermore, although *negative affect* was a significant predictor of *pragmatic quality*, *usefulness of content*, *perceived enjoyment* and *goodness* in the model of interaction experience with news sites in Study 2 (see Chapter 5, Section 5.3.4), it was the weakest predictor in the model with f^2 effect

sizes ranging from .02 to .04. Based on psychometric considerations and on the assumption that effect sizes this low may not reach significance on a sample of *N* = 85, *negative affect* was removed from Study 3.

6.3.3. Discriminant validity and descriptives of latent variables

In order to assess the discriminant validity of measures at the construct level, the square root of values of average variance extracted for each scale and each scale's correlation with all other measures are presented in Table 6.4. According to the Fornell-Larcker criterion (Fornell & Larcker, 1981), latent variables should share more variance with their indicators than with other latent variables; in other words, values of square root of average variance extracted should exceed scale inter-correlations. For discriminant validity at the item-level, see cross-loadings in Appendix 6.3. Descriptive statistics of latent variables[22] are presented in Table 6.5.

[22] Latent variable scores were calculated by running the PLS algorithm (see Chapter 4, Section 4.4.1).

Table 6.4. Coefficients of discriminant validity.

	1	2	3	4	5	6	7	8	9	10
1 Acumen CVPA	0.89									
2 Aesthetics	.11	0.76								
3 Adequacy of information	.04	**.59	0.90							
4 Beauty T1 (before use)	.22	**.63	**.32	1.00						
5 Beauty T2 (after use)	*.22	**.68	**.34	**.60	1.00					
6 Behavioural intention	*.27	**.52	**.35	**.38	**.47	0.99				
7 Goodness T1	*.24	**.69	**.53	**.68	**.54	**.52	1.00			
8 Goodness T2	.16	**.76	**.49	**.54	**.68	**.69	**.70	1.00		
9 Hedonic quality T1	*.27	**.63	**.50	**.75	**.50	**.45	**.72	**.59	0.79	
10 Hedonic quality T2	.14	**.87	**.51	**.63	**.73	**.54	**.65	**.81	**.72	0.83
11 Perceived disorientation	-.08	**-.51	**-.38	**-.41	**-.49	**-.39	**-.48	**-.52	**-.32	**-.42
12 Perceived enjoyment	*.26	**.67	**.48	**.53	**.62	**.53	**.65	**.80	**.64	**.76
13 Positive affect	**.32	**.64	**.45	**.45	**.64	**.47	**.49	**.63	**.52	**.67
14 Pragmatic quality T1	.18	**.46	**.50	**.40	*.26	**.31	**.63	**.46	**.43	**.34
15 Pragmatic quality T2	.12	**.65	**.45	**.49	**.59	**.46	**.62	**.66	**.48	**.63
16 Response CVPA	**.32	.17	*.25	.01	-.01	.14	.11	.10	.14	.10
17 Spatial ability	.18	-.02	.00	.05	.05	.20	.03	.10	.10	.01
18 Satisfaction	*.23	**.73	**.64	**.52	**.57	**.67	**.76	**.79	**.62	**.71
19 Trust	.13	.21	**.52	.16	.05	**.33	*.24	.10	**.30	.11
20 Usefulness of content	.22	**.56	**.60	**.35	**.37	**.55	**.47	**.59	**.44	**.52
21 Value CVPA	**.52	.09	.19	.04	.05	.05	.19	-.07	.21	.05

(Continued on next page)

Table 6.4 (continued).

	11	12	13	14	15	16	17	18	19	20	21
11 Perceived disorientation	0.76										
12 Perceived enjoyment	**-.53	0.93									
13 Positive affect	**-.33	**.75	0.74								
14 Pragmatic quality T1	**-.35	**.37	.22	0.83							
15 Pragmatic quality T2	**-.70	**.71	**.53	**.50	0.84						
16 Response CVPA	.05	.07	*.21	.12	-.01	0.84					
17 Spatial ability	-.18	.12	-.03	-.04	.10	-.01	1.00				
18 Satisfaction	**-.58	**.74	**.61	**.52	**.69	*.23	.13	0.92			
19 Trust	-.01	.07	.13	.19	.07	**.41	.04	**.34	0.87		
20 Usefulness of content	**-.34	**.51	**.57	**.33	**.44	.20	.18	**.67	**.34	0.79	
21 Value CVPA	.09	.01	*.22	.03	-.05	**.51	.02	.12	*.25	*.23	0.84

Note. Diagonal elements are square root of average variance extracted. Off-diagonal elements are correlation coefficients.

*$p < .05$. **$p < .01$.

Table 6.5. Descriptive statistics of latent variables.

Scale	Number of items	Mean	Standard deviation
Acumen CVPA	2	4.25	1.41
Aesthetics	8	4.40	1.19
Adequacy of information	3	5.05	1.30
Beauty T1 (before use)	1	4.39	1.42
Beauty T2 (after use)	1	4.54	1.36
Behavioural intention	2	5.16	2.15
Goodness T1	1	5.06	1.52
Goodness T2	1	5.44	1.52
Hedonic quality T1	4	4.61	1.16
Hedonic quality T2	4	4.81	1.25
Perceived disorientation	7	2.24	1.12
Perceived enjoyment	3	4.96	1.47
Positive affect	10	4.22	1.28
Pragmatic quality T1	3	4.91	1.21
Pragmatic quality T2	3	5.29	1.26
Response CVPA	3	4.55	1.57
Spatial ability	1	2.95	2.45
Satisfaction	3	4.89	1.34
Trust	4	5.37	1.05
Usefulness of content	3	5.65	1.00
Value CVPA	4	4.89	1.44

The coefficients of discriminant validity at the construct level in Table 6.4 indicate that *perceived aesthetics* shared more variability with *hedonic quality T2* (r = .87, R^2 = 76%) than the amount of variance it shared on average with its own items (*AVE* = 58%). However, coefficients of discriminant validity at the item level presented in Appendix 6.3 indicate that each classical- and expressive-aesthetics item loaded stronger on the *perceived aesthetics* scale than on *hedonic quality T2*, although four aesthetics items have cross-loadings greater than .70 on *hedonic quality T2*. *Perceived aesthetics* shared the equal amount of variance with its items and with *goodness T2* (58%), but analysis at the item level showed that each aesthetics item loaded higher on *perceived aesthetics* than on *goodness T2*; moreover, only one

aesthetics item loaded over .70 on *goodness T2*. *Hedonic quality T2* shared closely equal amount of variance with its items (*AVE* = 69%) and with *goodness T2* (R^2 = 66%). However, *hedonic quality* items loaded higher on *hedonic quality T2* than on *goodness T2*, although two hedonic quality items had cross-loadings larger than .70 on *goodness T2*. Because both *perceived aesthetics* and *hedonic quality T2* exhibited high composite reliability (.92 and .90, respectively) and analyses of discriminant validity at the item level supported their relative independence from each other and from *goodness T2*, these three measures were treated as separate in further analyses. The violation of the Fornell-Larcker criterion for discriminant validity was a result of high scale inter-correlations between *perceived aesthetics*, *hedonic quality T2* and *goodness T2* (ranging from .76 to .87).

6.4. Analysis

6.4.1. Between-group differences on interaction-experience measures

Study 1 revealed significant differences between novice and experienced users of a news site in *perceived enjoyment*, *aesthetics* (classical and expressive), *hedonic quality*, *beauty* and *goodness* judgements, while differences in *intensity of flow* and *pragmatic quality* were not significant (see Chapter 3, Section 3.6.1). Similar to Study 1, the BBC and NZH groups in Study 3 differed in the level of adoption (novice or experienced users), but in Study 3, the artefacts were different too for the two groups. Note that due to the design of the current study, differences between the experimental groups stemming from the level of adoption, and from differences in design and content of the sites cannot be separated. However, based on the findings of Study 1, it was expected that the NZH as a non-adopted site would receive lower ratings on interaction-experience scales than BBC as an adopted site.

Independent samples *t* tests were conducted on each interaction-experience measure to explore differences between the BBC and NZH groups. The results of the *t* tests are summarised in Table 6.6.

Table 6.6. Tests of difference in interaction experience between the experimental conditions.

Variable	BBC Mean (*SE*)	NZH Mean (*SE*)	Mean diff.	t^a	df^b	*p*	Effect size *(r)*
Pragmatic quality (T1)[c]	5.09 (.17)	4.74 (.20)	0.35	1.36	83	.196	0.15
Hedonic quality (T1)	5.07 (.15)	4.14 (.17)	0.93	3.99	83	***.000	0.40
Beauty (T1)	4.84 (.20)	3.93 (.22)	0.91	3.11	82.04	**.003	0.32
Goodness (T1)	5.56 (.20)	4.55 (.24)	1.01	3.23	83	**.002	0.33
Pragmatic quality (T2)	5.53 (.17)	5.03 (.21)	0.50	1.86	83	.067	0.20
Hedonic quality (T2)	5.17 (.17)	4.44 (.20)	0.73	2.80	83	**.006	0.29
Beauty (T2)	4.74 (.16)	4.33 (.24)	0.41	1.40	71.72	.167	0.16
Goodness (T2)	6.00 (.15)	4.86 (.27)	1.14	3.71	62.73	***.000	0.42
Perceived disorientation	2.00 (.15)	2.49 (.19)	-0.49	-2.04	83	*.044	0.22
Perceived enjoyment	5.35 (.18)	4.57 (.25)	0.78	2.50	83	*.014	0.26
Positive affect	4.46 (.20)	3.98 (.19)	0.48	1.76	83	.083	0.19
Perceived aesthetics	4.75 (.17)	4.05 (.18)	0.70	2.82	83	**.006	0.30
Usefulness of content	5.98 (.13)	5.30 (.16)	0.68	3.29	83	***.001	0.34
Adequacy of information	5.28 (.19)	4.82 (.20)	0.46	1.65	83	.103	0.18
Trust	5.71 (.14)	5.02 (.16)	0.69	3.14	83	**.002	0.33
Behavioural intention	6.41 (.13)	3.88 (.36)	2.53	6.60	51.94	***.000	0.68
Satisfaction	5.39 (.15)	4.37 (.23)	1.02	3.76	70.61	***.000	0.41

Note. All variables were measured on a 7-point scale.

[a]Bootstrap, $N = 5000$.

[b]Fractional degrees of freedom are presented where Levene's test for equality of variances was significant.

[c]T1: before use. T2: after use

*$p < .05$. **$p < .01$. ***$p < .001$.

With regards to outcome variables, BBC received significantly higher ratings on *behavioural intention*, *satisfaction*, *goodness* (both before and after use) and before-use *beauty*. BBC was not rated significantly higher on *beauty* than NZH after use (T2); however, BBC was rated significantly more aesthetic than NZH after use (T2). With regards to measures of components of interaction experience, BBC received significantly higher ratings of *perceived enjoyment*, *usefulness of content*, *trust* and

hedonic quality (both before and after use). Similar to the results of Study 1, ratings of *pragmatic quality* did not differ significantly in the two conditions. Moreover, ratings of *pragmatic quality* in Study 3 were not significantly different between the groups either before use (after presentation) or after use. Participants in both conditions were experienced users of news sites, including BBC; therefore, they were familiar with the general layout of news sites. The results suggest that participants did not perceive NZH, which they encountered for the first time, less usable than BBC. Furthermore, ratings of *adequacy of information* and *positive affect* did not differ significantly between the experimental conditions.

Although there was a significant difference in before-use *beauty* judgements in favour of BBC, it was not rated significantly more beautiful than NZH after use, which suggests a change in *beauty* judgements between before and after use. Previous research found judgements of *beauty* and perceptions of *hedonic quality* to be stable over time, whereas perceived usability and ratings of *goodness* were affected by experience (Hassenzahl, 2004). Participants completed the AttrakDiff2 questionnaire twice during the experiment: after a self-paced exposure of five screen-shots (before use) and after 10 minutes use of the news site (after use). To assess the effect of interaction with the news sites on perceptions of product attributes and quality judgements, paired-samples *t* tests were conducted between before and after use. The tests of stability of judgements of *pragmatic quality*, *hedonic quality*, *beauty* and *goodness* are presented in Table 6.7, separately for the BBC and NZH groups.

Table 6.7. Paired-samples *t*-tests for AttrakDiff2 in the two experimental conditions.

Group	Variable	T1 Mean (*SE*)	T2 Mean (*SE*)	Mean diff.	*t*	*df*	*p*	Effect size (*r*)
BBC	Pragmatic quality	5.09 (.17)	5.53 (.17)	-0.44	-2.95	42	**.005	0.41
	Hedonic quality	5.07 (.15)	5.17 (.17)	-0.10	-0.87	42	.390	0.13
	Beauty	4.84 (.20)	4.74 (.16)	0.10	0.52	42	.605	0.08
	Goodness	5.56 (.20)	6.00 (.15)	-0.44	-2.50	42	*.017	0.36
NZH	Pragmatic quality	4.74 (.20)	5.03 (.21)	-0.29	-1.33	41	.190	0.20
	Hedonic quality	4.14 (.17)	4.44 (.20)	-0.30	-1.87	41	.069	0.28
	Beauty	3.93 (.22)	4.33 (.25)	-0.40	-2.07	41	*.045	0.31
	Goodness	4.55 (.24)	4.86 (.27)	-0.31	-1.65	41	.108	0.25

For BBC, *hedonic quality* and *beauty* judgements proved to be stable; they did not differ significantly before use and after use. However, *pragmatic quality* and *goodness* ratings were significantly higher after use. These findings are in line with that of Hassenzahl (2004), who argues that beauty judgements are not strongly affected by experience and hedonic attributes can be derived from the appearance of products, whereas use experience may provide further information about usability. However, the pattern of results shows the opposite for NZH: *pragmatic quality* and *goodness* ratings proved to be stable, whereas *beauty* ratings were significantly higher after use, and the increase of *hedonic quality* ratings after use approached significance. In an experiment using an information-oriented Web site, van Schaik and Ling (2009) found that ratings of perceived aesthetics increased after the use of the site. Although the present study did not measure *perceived aesthetics* before use, the current findings indicate that use experience can provide further information about the appearance of a news site as a whole.

After-use *hedonic quality* and *goodness* were significantly different between the experimental groups. Two ANCOVAs were conducted to compare the BBC and NZH groups on after-use *hedonic quality* and *goodness*, with before-use *hedonic quality* and *goodness* as covariates, respectively. The ANCOVA testing the effect of

experimental group on after-use *hedonic quality* with before-use *hedonic quality* as covariate found the effect of experimental group non-significant, F (1, 81) = .004, p = .95, ε^2 = -.01. However, the effect of before-use *hedonic quality* as a covariate was significant, F (1, 81) = 68.52, p < .001, ε^2 = .33.[23] This finding suggests that the difference between the BBC and NZH groups in after-use *hedonic quality* was attributable to differences in before-use *hedonic quality* between the groups. The assumption of homogeneity of regression slope was checked using the interaction between experimental condition and before-use *hedonic quality*. The interaction effect was non-significant and explained a very small proportion of variance, F (1, 81) = .009, p = .92, ε^2 = -.01.

An ANCOVA was conducted to test the effect of experimental group on after-use *goodness* as dependent variable with before-use *goodness* as covariate. However, the assumption of homogeneity of regression slope was violated, F (1, 81) = 8.51, p < .01, ε^2 = .04. Although the assumption was violated, the effect size of the interaction effect of the experimental condition as independent variable and before-use *goodness* as covariate was between small and medium, according to the effect-size convention suggested by Cohen (1988) (.01 - small, .06 - medium and .14 - large)[24]. Therefore, the results of the ANCOVA are presented here, but they need to be interpreted with some caution. The effect of experimental group on after-use *goodness* was significant, F (1, 81) = 11.50, p = .001, ε^2 = .06. The effect of before-

[23] Epsilon squared is used here as a relatively unbiased measure of effect size. The formula for epsilon squared is provided by Keppel (1982).

[24] Moreover, in relation to checking the assumption of homogeneity of regression slope in ANCOVA, Clark-Carter (2010) suggests that researchers should not depend entirely on statistical significance, but should also assess the proportion of variance the interaction explains.

use *goodness* as a covariate was significant, F (1, 81) = 58.78, p < .001 , ε^2 = .34. Although both effects were significant, the higher effect size of the covariate suggests that differences in after-use *goodness* between the groups were predominantly attributable to differences in before-use *goodness* ratings.

On average, participants visited 22.51 pages (*SD* = 10.66) during the 10 minutes of browsing the news site. The average number of home-page revisits was 3.18 (*SD* = 2.55). There was no significant difference between number of pages visited by the BBC-user group (*M* = 22.51, *SE* = 1.66) and the NZH-user group (*M* = 22.50, *SE* = 1.63), *t* (83) = 0.01, *p* = .10. Similarly, the number of home-page revisits for the BBC group (*M* = 3.09, *SE* = 0.43) and the NZH group (*M* = 3.26, *SE* = 0.53) was not significantly different, *t* (83) = -0.30, *p* = .76. Although the NZH group reported higher level of disorientation than the BBC group, *perceived disorientation* was not correlated significantly with the number of home page revisits (BBC: *r* = -.04, *p* = .79; NZH: *r* = -.02, *p* = .90) and the number of pages visited (BBC: *r* = .18, *p* = .24; NZH: *r* = -.18, *p* = .25).

6.4.2. Testing the model of interaction experience with news sites

In order to test the model of interaction experience with news sites formulated in Chapter 5, the structural relations between artefact characteristics, components of interaction experience and outcome measures were drawn according to this model (Chapter 5, Section 5.3.4; see Figure 5.8). Note that the current study did not include measures of *perceived user-interface design* and *accessibility* among artefact characteristics, and *negative affect* was removed from the measurement model based on psychometric considerations. Also note that after-use (T2) *pragmatic quality*, *hedonic quality*, *beauty* and *goodness* are used in the testing of the model.

Hypotheses related to testing the model of interaction experience with news sites are presented in Figure 6.2. To test the relationships between artefact characteristics and components of interaction experience, the following hypotheses are proposed, based on the findings of Study 2 and the theoretical considerations related to each hypothesis, presented in Chapter 5:

Hypothesis 1: *perceived disorientation* is an antecedent of *pragmatic quality*.

Hypothesis 2: *adequacy of information* is an antecedent of *usefulness of content*.

Hypothesis 3: *perceived aesthetics* is an antecedent of *positive affect*.

Hypothesis 4: *perceived aesthetics* is an antecedent of *hedonic quality*.

In the model of interaction experience with news sites in Study 2, the effect of *perceived aesthetics* on *perceived enjoyment* was fully mediated through *hedonic quality*. Therefore, the following hypotheses are proposed:

Hypothesis 5: *perceived aesthetics* is not a direct antecedent of *perceived enjoyment*.

Hypothesis 6: *hedonic quality* is an antecedent of *perceived enjoyment*.

In Study 2, when artefact characteristics were excluded from the analysis, *positive affect* was significantly connected to all measures of perceptions of instrumental and non-instrumental qualities in the model of interaction experience with news sites

(Chapter 5, Section 5.3.4; see Table 5.7)[25]. *Positive affect* was a determinant of *beauty*, but not *goodness*. Therefore, the following hypotheses are proposed:

Hypothesis 7: *positive affect* is an antecedent of *hedonic quality*.

Hypothesis 8: *positive affect* is an antecedent of *perceived enjoyment*.

Hypothesis 9: *positive affect* is an antecedent of *pragmatic quality*.

Hypothesis 10: *positive affect* is an antecedent of *usefulness of content*.

Hypothesis 11: *positive affect* is a determinant of *beauty*.

Hypothesis 12: *positive affect* is not a determinant of *goodness*.

Based on the results of hypothesis tests in Study 2, the following hypotheses are proposed for the remainder of the connections between components of interaction experience and interaction outcomes:

Hypothesis 13: *hedonic quality* is a determinant of *beauty*.

Hypothesis 14: *hedonic quality* is a determinant of *goodness*.

Hypothesis 15: *pragmatic quality* is a determinant of *goodness*.

Hypothesis 16: *pragmatic quality* is not a determinant of *beauty*.

Hypothesis 17: *perceived enjoyment* is a determinant of *behavioural intention*.

Hypothesis 18: *usefulness of content* is a determinant of *behavioural intention*.

---

[25] Note that the direction of connections between positive affect and components of interaction experience may be reverse. For a detailed discussion, see Chapter 5, Section 5.3.4).

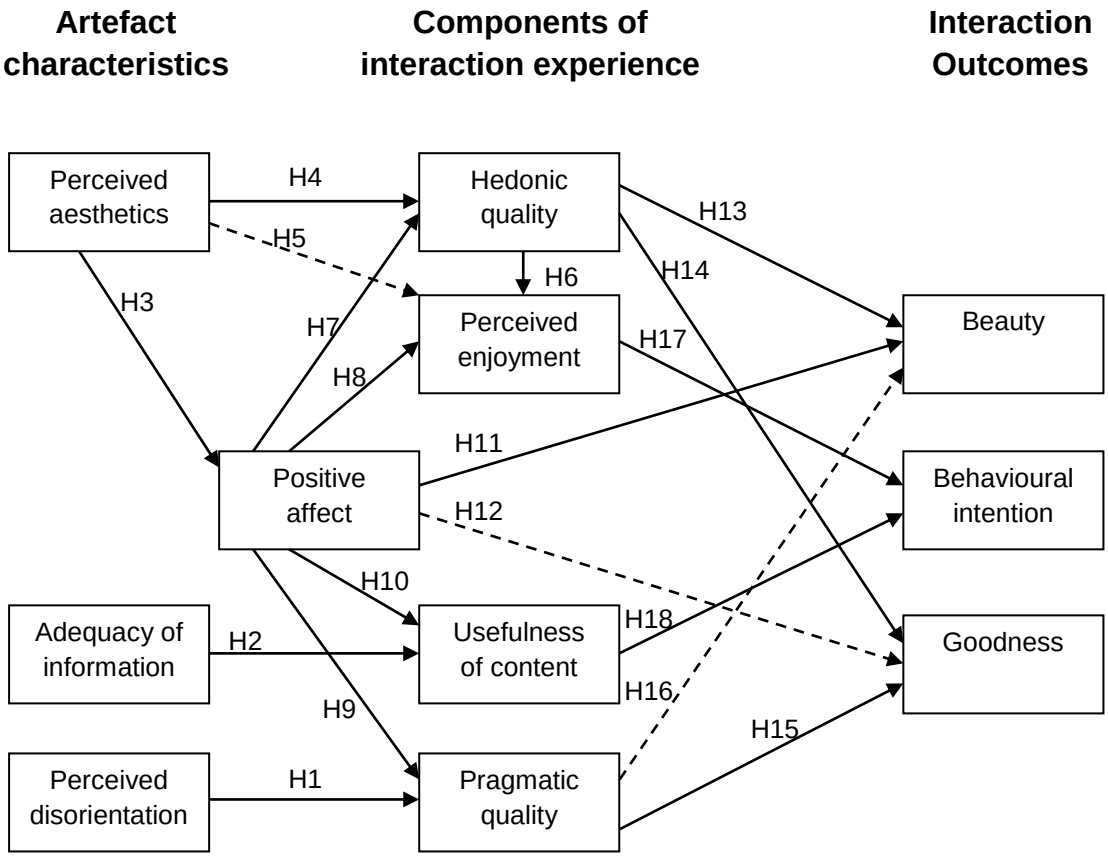


Figure 6.2. Summary of hypotheses related to the model of interaction experience with news sites.

The tests of Hypotheses 1-18 are presented in Table 6.8. All hypotheses except Hypothesis 16 were supported. Contrary to Hypothesis 16, *pragmatic quality* was found to be a significant predictor of *beauty*, however, it had the lowest path coefficient in the model (β = .18). In Hassenzahl's (2003, 2004) model of interaction experience, *pragmatic quality* is not a determinant of *beauty*. Similarly, van Schaik and Ling (2011) found that while *goodness* judgements of a Web-based encyclopaedia were influenced by both *hedonic quality* and *pragmatic quality*, *beauty* judgements were only influenced by ratings of *hedonic quality*. Therefore, an inference perspective from overall quality judgements (*goodness* and *beauty*) to

product to perceived product qualities (pragmatic and hedonic) was adopted to examine the relationship between *pragmatic quality* and *beauty* (see Hassenzahl & Monk, 2010 and van Schaik et al., under review).

Table 6.8. Summary of tests of hypotheses related to the model of interaction experience with news sites.

	Predictor variable	Outcome variable	β	t^a	Supported (Yes/No)
H_1	Perceived disorientation	Pragmatic quality	-.60	***9.96	Yes
H_2	Adequacy of information	Usefulness of content	.44	***3.72	Yes
H_3	Perceived aesthetics	Positive affect	.65	***10.12	Yes
H_4	Perceived aesthetics	Hedonic quality	.76	***12.33	Yes
H_5	Perceived aesthetics	Perceived enjoyment	-.12	0.83	Yes
H_6	Hedonic quality	Perceived enjoyment	.57	***3.38	Yes
H_7	Positive affect	Hedonic quality	.18	*2.57	Yes
H_8	Positive affect	Perceived enjoyment	.44	***4.55	Yes
H_9	Positive affect	Pragmatic quality	.34	***4.60	Yes
H_{10}	Positive affect	Usefulness of content	.36	**3.27	Yes
H_{11}	Positive affect	Beauty	.24	*2.53	Yes
H_{12}	Positive affect	Goodness	.09	1.18	Yes
H_{13}	Hedonic quality	Beauty	.46	***4.46	Yes
H_{14}	Hedonic quality	Goodness	.60	***5.34	Yes
H_{15}	Pragmatic quality	Goodness	.24	*2.01	Yes
H_{16}	Pragmatic quality	Beauty	.18	*2.03	No
H_{17}	Perceived enjoyment	Behavioural intention	.35	***3.53	Yes
H_{18}	Usefulness of content	Behavioural intention	.38	***3.47	Yes

[a]Bootstrap, $N = 5000$.

*$p < .05$. **$p < .01$. ***$p < .001$.

6.4.3. The relationship between pragmatic quality and beauty

There is an abundance of reports in the human-computer interaction literature about the relationship between beauty and perceived usability[26]. In a review of 17 papers, Hassenzahl and Monk (2010) found that Pearson correlations between beauty[27] and perceived usability ranged from .03 to .92, with a median of .61. In the current study,

[26] Perceived usability is conceptualised in the model as pragmatic quality and the two terms are used interchangeably.

[27] Beauty was usually conceptualised as classical aesthetics in the reviewed studies.

the correlation between *beauty* and *pragmatic quality* before use (after the presentation of the screenshots) was $r = .40$, $p < .01$ (BBC group: $r = .33$, $p < .05$; NZH group: $r = .41$, $p < .01$), and $r = .59$, $p < .01$ after interacting with the sites (BBC group: $r = .49$, $p < .05$; NZH group: $r = .63$, $p < .01$).

The model of interaction experience with news sites is causative, where interaction outcomes (quality judgements and intention to use) are consequences of components of interaction experience (perceived product character, enjoyment, usefulness and affect). However, causative models are limited in a way that they treat evaluative outcome variables exclusively as consequences of perceived qualities and ignore the possibility of general-to-specific inference processes in judgements (Hassenzahl & Monk, 2010; van Schaik et al., under review).

Based on Lindgaard et al. (2006), Hassenzahl and Monk (2010) argue that judgements of beauty play an important role in inferring characteristics of artefacts (such as usability) that are unavailable at the point of judgement, because aesthetic information is ubiquitous and immediately available (without hands-on experience). Hassenzahl and Monk (2010) found support for the notion that *pragmatic quality* is inferred indirectly from *beauty* through *goodness*. This process of indirect inference is referred to as *evaluative consistency* (Lingle & Ostrom, 1979), which assumes that individuals infer a general evaluation of goodness from perceptions of beauty, and unavailable attributes, such as usability, are inferred from this goodness evaluation. The authors also found that *hedonic quality* was directly inferred from *beauty*. This process of direct inference of an unavailable attribute from an available one is referred to as *probabilistic consistency* (Ford & Smith, 1987).

Because inference processes in evaluative judgements become especially applicable in the absence of hands-on experience (van Schaik et al., under review), it is assumed that before use the NZH group were the most reliant on inferences from *beauty* in the current study, because the only contact they had had at that point with the artefact was the self-paced exposure of five non-interactive screenshots of different pages of the site. Although users of BBC in the experiment all had previous experience of using the site, inferential processes are expected to play a part in formulating judgements of product qualities and evaluation before use. In a series of experiments, van Schaik and colleagues (under review) found support for the inference model both after self-paced presentation of screenshots and after actual use, both in action mode (exploration as task) and goal mode (information-retrieval tasks), and both with a Web site that was novel to users and with a Web site the users had used before. Therefore, the inference-perspective model was expected to be supported both after presentation and after use, and both for NZH (novel artefact) and BBC (adopted artefact). General-to-specific inference-perspective models were specified separately for NZH and BBC in order to examine between-group differences.

In a study using four different samples and a wide range of Web sites, Hassenzahl and Monk (2010) found a generally low correlation between *hedonic quality* and *pragmatic quality* (average $r = .23$), supporting the independence of the two constructs. In the three studies of van Schaik et al. (under review), the average correlation between *hedonic quality* and *pragmatic quality* was .41 before use and .38 after use. The correlation between *pragmatic quality* and *hedonic quality* in the current study was .43 before use ($p < .01$) and .63 after use ($p < .01$) in the pooled sample, indicating a large amount of variance shared between the two constructs. In

particular, the correlation between the two constructs was .47 ($p < .01$) before use and .65 ($p < .01$) after use in the NZH condition, and .32 ($p < .05$) before use and .53 ($p < .01$) after use in the BBC condition. Therefore, based on van Schaik and colleagues (under review) and Hassenzahl and Monk (2010), a path from *pragmatic quality* to *hedonic quality* was drawn in the models to test the independence of the two constructs after controlling for their relationships with *beauty* and *goodness*. Inference-perspective analyses of NZH before and after use are presented in Figure 6.3, and the same analyses of BBC are presented in Figure 6.4.

Note that although PLS requires a smaller sample size than covariance-based structural equation modelling (Chin, 2010), it still requires an appropriate sample size to detect effects with adequate power. The sample sizes of $n = 43$ for the BBC group and $n = 42$ for the NZH group reach the minimum recommended by Chin (1998)[28]. However, rules of thumb cannot be fully trusted to accurately approximate appropriate sample sizes and an adequate sample size is still required to detect smaller effect sizes (Marcoulides & Saunders, 2006). As a consequence, smaller effect sizes may remain undetected when sample size is close to the minimum recommendation based on rules of thumb.

[28] According to Chin's (1998) rule of thumb, the minimum sample size should be ten times the number of the larger of two numbers: (a) the number of indicators of the latent variable with the largest number of indicators, and (b) the number of paths of the latent variable with the largest number of paths directed to it. In the present case, the minimum sample size is 4 x 10 = 40 (based on [a]).

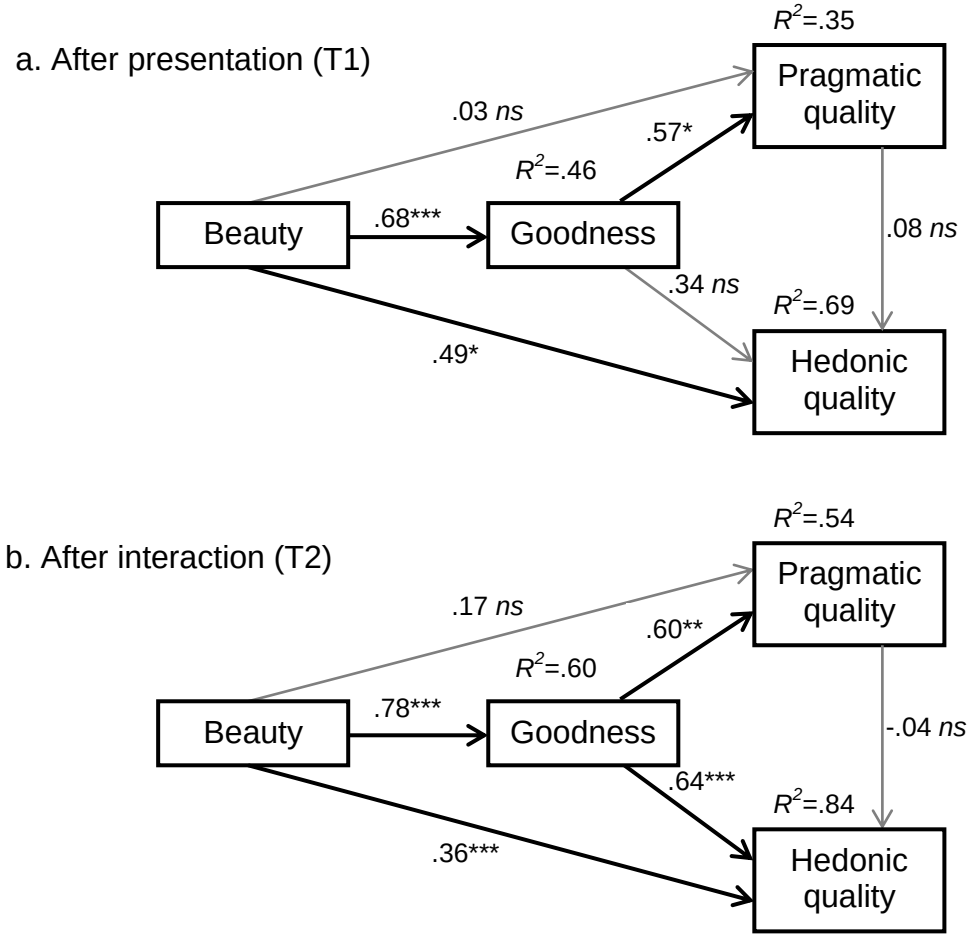


Figure 6.3. Inference model of the NZH group.

Figure 6.3 shows that the direct path from *beauty* to *pragmatic quality* was non-significant both before use and after use for NZH, but the indirect effect of *beauty* on *pragmatic quality* was .39 (z = 2.02, p < .05)[29] before use and .47 (z = 3.07, p < .01) after use, supporting evaluative consistency. On the other hand, *beauty* was a direct determinant of *hedonic quality* both before and after use, supporting probabilistic consistency; however, the effect of *beauty* on *hedonic quality* was partially mediated

[29] Significance of indirect paths was calculated using the Sobel test (see MacKinnon et al., 2002).

through *goodness* after use, with the indirect effect larger than the direct effect (β = .50, z = 5.31, p < .001, proportion of the total effect that is mediated (mediated effect or *ME* for short) = .58) [30]. The inference perspective model was supported for NZH at both times of measurement, indicating the lack of direct relationship between ratings of *beauty* and perceived usability. In general, the strength of correlation and connections between constructs was stronger, and the prediction of latent variables was better after use than at before use, indicating a stronger relationship between constructs after use for the NZH condition.

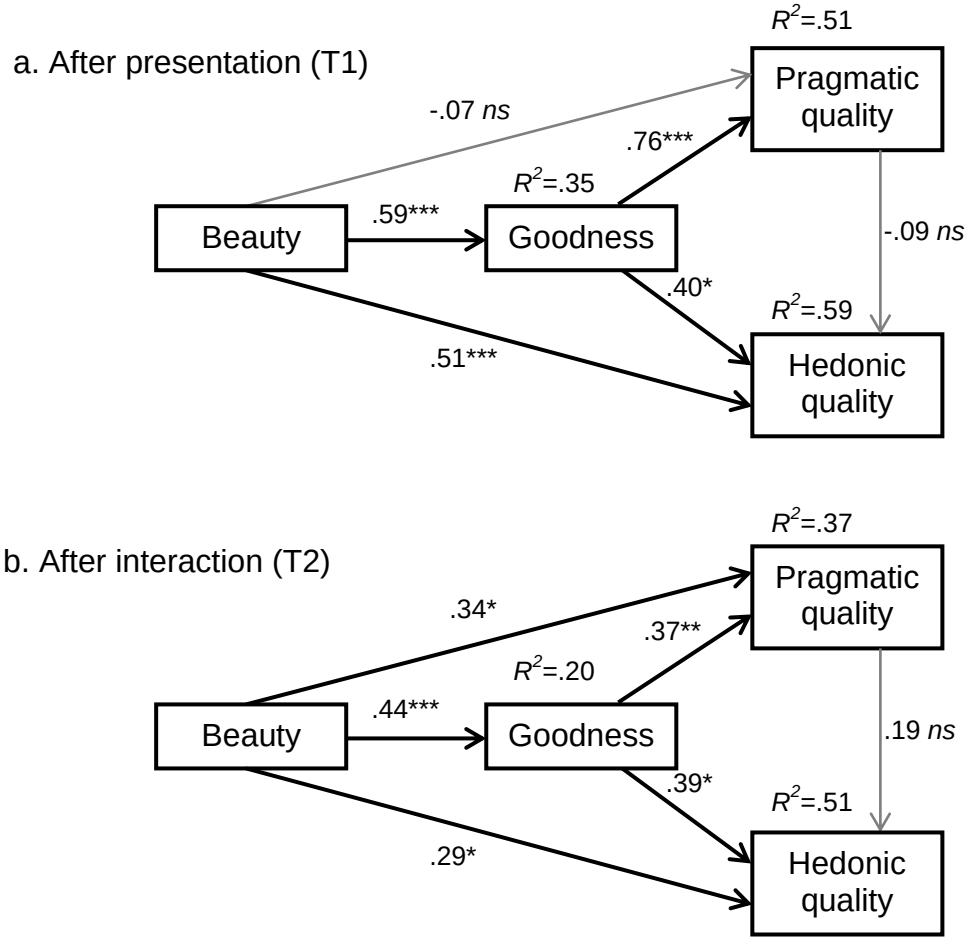


Figure 6.4. Inference model of the BBC group.

[30] Values of mediated effect are presented where the direct and indirect effects are both significant.

Analysis of the BBC group revealed that the inference-perspective model was supported before use, but not after use. Before use, the direct effect of *beauty* on *pragmatic quality* was non-significant (indirect effect = .45, $z = 4.38$, $p < .001$), supporting evaluative consistency. The direct effect of *beauty* on *hedonic quality* was significant before use, supporting probabilistic consistency, but its indirect effect through *goodness* was significant too ($\beta = .24$, $z = 1.96$, $p = .05$, $ME = .32$). After use, there was a significant direct connection between *beauty* and *pragmatic quality* after controlling for *goodness* (indirect effect = .16, $z = 2.14$, $p < .05$, $ME = .32$), and over two thirds of the total effect was accounted for by the direct effect. The direct connection between *beauty* and *hedonic quality* was significant ($\beta = .29$, $t = 1.97$, $p = .050$; indirect effect = .17, $z = 2.32$, $p < .05$, $ME = .37$).

The significant and nearly medium effect-size ($\Delta R^2 = .09$, $f^2 = .14$) direct path between *beauty* and *pragmatic quality* in the BBC group after use suggests a direct relationship between perceived usability and the judgement of appearance of the site, and contradicts the findings of Hassenzahl and Monk (2010) and van Schaik and colleagues (under review). However, directional assumptions cannot be drawn based on regression lines and it may be that a direct relationship is observed because an important mediator of the direct effect is not addressed in this model. Nevertheless, it is noteworthy that the total effect (the sum of the direct and indirect effects) of *beauty* on *pragmatic quality* (.50) exceeds its total effect on *hedonic quality* (.46). Contrary to the results of the NZH condition, the prediction accuracy of latent variables in terms of determination coefficients was lower after use than before use. A test of the inference model after use on a pooled sample revealed that the results of the BBC group biased the model to such an extent that the direct effect between *beauty* and *pragmatic quality* remained significant ($\beta = .26$, $t = 3.10$, $p < .01$).

Therefore, it can be concluded that the results of the BBC group are responsible for the significant, although low-impact, path from *pragmatic quality* to *beauty* in the model of interaction experience with news sites (Hypothesis 16). In addition, the relationship between *pragmatic quality* and *hedonic quality* became non-significant when controlling for the effects of *beauty* and *goodness* in both groups, both before and after use, supporting the independence of the two dimensions of perceived product attributes despite the high correlation observed between them.

6.4.4. Additions to the model

6.4.4.1. Trust

Based on a review of literature and in-depth interviews with news-media experts, Chen and Corkindale (2008) identified trust as one of the main drivers in the adoption of online news services. In their technology acceptance model for online news, trust serves as a determinant of *perceived usefulness* and a direct determinant of *behavioural intention* (see Chapter 2, Figure 2.4). There is empirical support for trust as a direct determinant of behavioural intention in the technology acceptance model applied to the domain of online retail (e.g., Gefen et al., 2003; Pavlou & Gefen, 2004). It is therefore expected that the addition of trust to the model would increase the prediction of *behavioural intention*. However, no psychometric instrument addressing trust was selected for the measurement model of Study 2, because the available measures of online trust either typically apply to the contexts of online retail and e-commerce, or aim to measure the coverage of particular news events and news topics instead of the perceived trustworthiness of a particular news provider or site (see Chapter 4, Section 4.2.6.1). Study 2 revealed significant differences in model parameters between users of the BBC site and other news sites. In particular,

predictions of the outcome measures were weaker for the BBC site than other news sites. In Study 3, *trust* was therefore added to the model to examine its effect on, and improve the prediction of, *behavioural intention*.

In correspondence with trust in journalistic assessment as a dimension of trust in news media (Kohring & Matthes, 2007), the four items were worded to gauge *trust* as the perceived truthfulness and reliability of information presented on the site, and trust in the competence of journalists providing information to the site. Item 3 (see Appendix 6.1) was adopted from Lavie and Tractinsky (2004), and it was originally used to measure service quality. In the measurement model, the scale exhibited satisfactory internal consistency (*CR* = .92) and retained 75% of variance from its items (*AVE* = .75). All item loadings were ≥ .80 (see Table 6.3). The scale exhibited satisfactory discriminant validity on the construct level (see Table 6.4), and the cross-loadings showed good discriminant validity at the item level (see Appendix 6.3).

To test if *trust* is a determinant of *behavioural intention*, *trust* was added to the structural model tested in Section 6.4.2. Based on Chen and Corkindale (2008), it is proposed that *trust* is an antecedent of *usefulness of content* (a proxy for *perceived usefulness* in the current model). There is no theoretical reason to propose that *trust* is connected to *beauty* and *goodness* judgements. In fact, it is reasonable to assume that overall *beauty* judgements of news sites are independent from the extent to which users deem a news site trustworthy. *Goodness* judgements would be more plausible to be affected by *trust* (however, Table 6.4 shows that there is no significant correlation between *goodness T2* and *trust*). Therefore, the following hypotheses are proposed:

Hypothesis 19: *trust* is a determinant of *behavioural intention*.

Hypothesis 20: *trust* is an antecedent of *usefulness of content*.

Hypothesis 21: *trust* is not a determinant of *beauty*.

Hypothesis 22: *trust* is a determinant of *goodness*.

The four paths above were added to the model, then the PLS and bootstrapping (N = 5000) algorithms were run. *Trust* was a significant predictor of *behavioural intention* (β = .22, t = 2.43, p < .05), but not of *beauty* (β = -.04, t = 0.62, *ns*) and *goodness* (β = .01, t = 0.08, *ns*). The hypothesis proposing that *trust* is an antecedent of *usefulness of content* was not supported (β = .11, t = 1.19, *ns*). Because *trust* was a significant direct predictor of *behavioural intention* (sr^2 = .04, f^2 = .07), controlling for *usefulness of content* and *perceived enjoyment*, it was retained in the model among the perception of instrumental qualities, as a determinant of *behavioural intention*.

6.4.4.2. Satisfaction

A measure of satisfaction was adopted as an additional outcome measure of experience, because it is a frequently used concept in expressing user's attitude about artefacts among design practitioners (Cho et al., 2011). As discussed in Chapter 2, the judgement of satisfaction involves an expectation, where one's experience of interaction with a particular artefact is judged in comparison to an initial position or standard. User-satisfaction or consumer-satisfaction is a main concept in e-commerce and marketing research (Ha, 2006); it is an important antecedent of loyalty to providers of services (Anderson & Srinivasan, 2003) and it is often used in relation to providing managerial guidance.

The 3-item satisfaction scale by Martensen & Grønholdt (2003) was adopted for the study. In compliance with different measurement practices in customer-satisfaction research (Ryan et al., 1995), the questionnaire contained an item each regarding (a) the overall level of satisfaction, (b) a judgement related to the extent to which a particular artefact meets one's expectations and (c) how close the artefact is to one's ideal (see Appendix 6.1 for item wording). Furthermore, the original scale was successfully applied by Martensen and Grønholdt in their study using PLS path modelling. In the measurement model, the *satisfaction* scale exhibited high internal consistency (*CR* = .94) and the scale retained 84% of variance from its three indicators. Discriminant validity was satisfactory, both at the construct and at the item level. Several items showed large cross-loadings on other scales, but the loadings exceeded the cross-loadings in every case (see Appendix 6.3).

Table 6.4 shows that *satisfaction* was positively and significantly correlated to all outcome measures[31], all measures of perceptions of instrumental and non-instrumental qualities, and *positive affect*. It is reasonable to assume that higher levels of perceptions of instrumental and non-instrumental qualities, and *positive affect* experienced during the interaction lead to higher levels of *satisfaction* in users of news sites. For example, Flavián and colleagues (2006) found that the usability of a Web site influences satisfaction[32], which in turn leads to greater loyalty to the site, and higher levels of trust were associated with higher levels of satisfaction. In order

[31] In an adaptation of TAM to Web technologies (Cho et al., 2009), user-satisfaction is included as a determinant of behavioural intention (see Chapter 2, Figure 2.5).

[32] Notably, along with effectiveness and efficiency of goal attainment, satisfaction is also part of the ISO 9241 definition of usability (see Chapter 2).

to explore the relationships between each of the predictors of outcome variables and *satisfaction*, the following hypotheses are proposed:

Hypothesis 23: *pragmatic quality* is a determinant of *satisfaction*.

Hypothesis 24: *hedonic quality* is a determinant of *satisfaction*.

Hypothesis 25: *usefulness of content* is a determinant of *satisfaction*.

Hypothesis 26: *perceived enjoyment* is a determinant of *satisfaction*.

Hypothesis 27: *positive affect* is a determinant of *satisfaction*.

Hypothesis 28: *trust* is a determinant of *satisfaction*.

Satisfaction was added to the path model and the connections presented in hypotheses 23-28 were drawn. The hypothesis tests related to *satisfaction* are presented in Table 6.9.

Table 6.9. Summary of tests of hypotheses related to predicting *satisfaction*.

	Predictor variable	β	t^a	Supported
H_{23}	Pragmatic quality	.23	*1.96	Yes
H_{24}	Hedonic quality	.19	1.69	No
H_{25}	Usefulness of content	.28	**2.70	Yes
H_{26}	Perceived enjoyment	.34	**2.71	Yes
H_{27}	Positive affect	-.08	0.86	No
H_{28}	Trust	.19	**2.96	Yes

[a]Bootstrap, N = 5000.

*p < .05, **p < .01, ***p < .001.

In accordance with the findings of Flavián and colleagues (2006), *pragmatic quality* (as a measure of user-perceived usability) and *trust* were positively and significantly related to *satisfaction*. *Usefulness of content* and *perceived enjoyment* were also significant predictors of *satisfaction*. Contrary to expectations, the effects of *hedonic*

quality and *positive affect* were not significant on *satisfaction*. However, both *hedonic quality* and *positive affect* serve as antecedents of *perceived enjoyment* in the model (see Figure 6.2 and Table 6.8); in fact, 68% of variance in *perceived enjoyment* is explained by these two variables (see Table 6.10 in the next section). It may be that the effects of *hedonic quality* and *positive affect* on *satisfaction* are fully mediated by *perceived enjoyment*. To test this assumption, *perceived enjoyment* was temporarily removed from the model, direct paths were drawn from *hedonic quality* and *positive affect* to *satisfaction*, and the path between *positive affect* and *hedonic quality* was removed to prevent any further partial mediating effects. As a result, the effect of *hedonic quality* became significant ($\beta = .30$, $t = 2.40$, $p < .05$), but the effect of *positive affect* remained non-significant ($\beta = .05$, $t = 0.64$, *ns*). Therefore, there is support for the idea that *hedonic quality* is a predictor of *satisfaction*, but its effect is mediated by *perceived enjoyment*. However, the path coefficient of *hedonic quality* ($\beta = .30$) was markedly lower than that of *perceived enjoyment*, when all non-significant paths were removed ($\beta = .40$, see Table 6.10 in the next section), and the prediction of *satisfaction* was slightly better with the inclusion of *perceived enjoyment* as its predictor ($R^2 = .74$) instead of *hedonic quality* ($R^2 = .72$). Therefore, *perceived enjoyment* was selected from among variables of perceptions of non-instrumental qualities to predict *satisfaction* in the model.

6.4.5. Model parameters

Following the extension of the model with *trust* and *satisfaction*, model parameters were calculated to describe all significant paths, and to assess the model's predictive power in terms coefficients of determination (R^2) estimated for each endogenous variable. Model parameters are presented in Table 6.10.

Table 6.10. Model parameters of the extended model.

Prediction of variables of interaction-experience components					
Target variable (number of predictors)	Predictor variable	β	t^a	sr^2	Unique effect size (f^2)[b]
Pragmatic quality (2)	Perceived disorientation	-.60	***9.89	0.32	0.80
R^2 = .60	Positive affect	.34	***4.62	0.10	0.26
Usefulness of content (2)	Adequacy of information	.44	***3.64	0.15	0.27
R^2 = .46	Positive affect	.35	**3.14	0.10	0.19
Perceived enjoyment (2)	Hedonic quality	.48	***4.76	0.13	0.39
R^2 = .68	Positive affect	.43	***4.40	0.10	0.31
Hedonic quality (2)	Perceived aesthetics	.76	***12.35	0.33	1.58
R^2 = .79	Positive affect	.18	**2.68	0.02	0.09
Positive affect (1)	Perceived aesthetics	.65	***10.06	R^2 = .42	0.72
Prediction of variables of interaction outcomes					
Target variable (number of predictors)	Predictor variable	β	t^a	sr^2	Unique effect size (f^2)[b]
Goodness (2)	Pragmatic quality	.26	*2.26	0.04	0.13
R^2 = .70	Hedonic quality	.65	***6.29	0.26	0.86
Behavioural intention (3)	Usefulness of content	.29	*2.47	0.06	0.10
R^2 = .43	Trust	.21	*2.37	0.04	0.07
	Perceived enjoyment	.24	***3.40	0.11	0.19
Satisfaction (4)	Pragmatic quality	.26	*2.12	0.03	0.13
R^2 = .74	Usefulness of content	.29	**2.89	0.05	0.21
	Trust	.19	**2.93	0.03	0.12
	Perceived enjoyment	.40	***3.47	0.08	0.29
Beauty (3)	Hedonic quality	.46	***4.55	0.09	0.23
R^2 = .59	Pragmatic quality	.18	*2.04	0.02	0.05
	Positive affect	.24	*2.55	0.03	0.08

[a]Bootstrap, N = 5000.

[b]f^2: 0.02 - 0.14 small, 0.15 - 0.34 medium and 0.35 - large.

*p < .05. **p < .01. ***p < .001.

Presumably due to differences in the design and level of control in data collection between the two studies, Study 3 yielded better predictions for each latent variable than Study 2, except for *usefulness of content*; however, *accessibility*, which was a predictor of *usefulness of content* in Study 2, was not included in Study 3. Most notably, the determination coefficients of outcome measures increased markedly: from 40% to 59% for *beauty*, from 47% to 70% for *goodness*, and from 36% to 43% for *behavioural intention*. *Trust* was added as a predictor of *behavioural intention* in Study 3. *Goodness* was predicted using the same variables in both studies, except

negative affect was excluded from the measurement model of Study 3. However, *negative affect* had the lowest path coefficient and effect size in Study 2 and its contribution to the prediction of *goodness* was minimal. Nevertheless, the prediction of *goodness* showed the largest increase among the outcome measures (23%). Although in Hassenzahl's (2003, 2004) user-experience model *pragmatic quality* is a determinant of *goodness*, but not of *beauty*, it was a significant predictor of *beauty* in Study 3. However, its path coefficient and effect size was the smallest in the entire model.

Satisfaction was introduced in the model in Study 3 and had the highest amount of explained variance among the outcome variables, with *perceived enjoyment*, *usefulness of content*, *pragmatic quality* and *trust* as its significant predictors. Because *perceived enjoyment* was (partly) predicted from *hedonic quality*, the direct effect of *hedonic quality* was also tested on *satisfaction* in the absence of *perceived enjoyment*. The direct effect of *hedonic quality* on *satisfaction* ($\beta = .33$, $t = 2.59$, $p < .01$) became non-significant in the presence of *perceived enjoyment* as a mediator ($\beta = .18$, $t = 1.55$, *ns*), supporting full mediation. Furthermore, *perceived enjoyment* had a stronger impact on *satisfaction* than *hedonic quality* (i.e., it had higher path coefficient and explained more variance). These results can be interpreted as follows: *hedonic quality*, the extent to which an artefact is perceived to be capable of promoting high-quality experience, promotes enjoyment during the interaction, which in turn promotes higher levels of *satisfaction* with the artefact.

There was a high positive correlation between after-use *hedonic quality* and *perceived aesthetics* ($r = .87$, $p < .01$). In fact, as discussed in the evaluation of the measurement model in Section 6.3.3, the scales after-use *hedonic quality* and

perceived aesthetics shared more variance with each other (76%) than with their predictors (69% and 58%, respectively)[33]. These findings indicate that the relationship between *perceived aesthetics* and *hedonic quality* need to be examined; particularly, are the two constructs different enough to justify the inclusion of both in the model.

The original two-dimensional perceived aesthetics scale (Lavie & Tractinsky, 2004) was designed to measure aesthetic attributes of interactive products. The *perceived aesthetics* scale in the current study is a one-dimensional combination of the items of an abridged version of this scale. According to Hassenzahl (2003, 2004), hedonic quality is a higher-order product quality perception that encompasses attributes of a particular product that promote pleasurable interactions with the product by fulfilling human needs (e.g., competence and stimulation). Aesthetic attributes may be considered as a part of pleasure-producing attributes that together are referred to as *hedonic quality*. Examination of the measurement items of the two scales supports this notion. The positive endpoints of the four semantic differentials used to measure *hedonic quality* are *captivating*, *stylish*, *premium* and *creative*. The four *expressive aesthetics* items are *original*, *sophisticated*, *spectacular* and *creative* (see Appendix 6.1). Note that the item *creative* is present in both scales, and the rest of the items of *hedonic quality* may be thought of as describing aesthetics too. However, the concept of hedonic quality is broader in scope than aesthetics in terms of

[33] As presented in Table 6.2, the *AVE* value of the aesthetics scale is lower than that of the subscales separately, which is a result of the combination of two originally separate, although correlated measurement dimensions. Lowered levels of *AVE* may lead to issues of discriminant validity when scale inter-correlations are high. For example, the square root of the *AVE* of the *expressive aesthetics* subscale was .87 (in contrast to .76 for the combined scale), which equalled the correlation between the combined *perceived aesthetics* scale and *hedonic quality T2* presented in Table 6.4.

summarising product attributes. Hedonic attributes are thought to describe product characteristics that are connected to be-goals, such as pleasure derived from the interaction, as opposed to do-goals (Hassenzahl & Roto, 2007), and hedonic attributes are theoretically connected to the fulfilment of general human needs (Hassenzahl et al., 2010). The high correlation between the two scales indicates that, at least in the current context, a high proportion of variance in hedonic attributes is attributable to aesthetic attributes. *Perceived aesthetics* was included in the model as part of perceived artefact characteristics, because it can be more directly connected to designable product characteristics than the more abstract construct of *hedonic quality*, which is considered in the model as a component of interaction experience. PLS path models were created to test if *beauty* as an outcome variable can be predicted more accurately from *hedonic quality* than from *perceived aesthetics*.

In a simplified path model, 48% of variance in *beauty* was predicted from *perceived aesthetics* alone. However, *hedonic quality* alone explained 54% of variance in *beauty*, and together with *positive affect*, explained variance rose to 58%. Moreover, mediation analysis in the full model revealed that the effect of *perceived aesthetics* on *beauty* is fully mediated through *positive affect* and *hedonic quality* (direct effect: $\beta = .10$, $t = 0.58$, $p > .05$). In summary, the simplified path models showed that *hedonic quality* provided a better prediction of *beauty* than *perceived aesthetics*, and the effect of *perceived aesthetics* on *beauty* was mediated through *hedonic quality* and *positive affect*. Although *hedonic quality* was a stronger predictor of *beauty*, *perceived aesthetic* can be more directly interpreted in relation to product design. Therefore, although the two scales were found to be very similar and highly overlapping in variance, the inclusion of both scales in the model is supported.

6.4.6. Differences in model parameters between the experimental conditions

Separate analysis of BBC users and users of other news sites in Study 2 (Chapter 5, Section 5.3.5) revealed differences in the prediction of outcome variables in the model. Participants in Study 2 were using adopted news sites, in other words, sites that they had been using before the study; therefore, differences in prediction of outcome variables could be attributed to differences between the sites rather than different stages of adoption. In Study 3, level of adoption and the artefact were different at the same time between the experimental groups. Separate analyses for the BBC and NZH groups were conducted to explore the differences between the experimental conditions and to test the robustness of the model. The results of between-group comparisons are presented in Table 6.11. The significance of differences between the path estimates was calculated using *t* tests, based on the differences between β-weights of the groups and the standard errors of the path estimates for each group obtained through bootstrapping resampling procedures (for calculating the pooled standard error, see Chin, 2000)[34].

[34] Although a non-parametric, permutation-based approach to group comparison in PLS was introduced recently, software support for this method is lacking, and a simulation study (Chin & Dibbern, 2010) indicates that it would not have the advantage of higher statistical power for the sample size and observed β differences in the present study.

Table 6.11. Differences of standardised path coefficients between the BBC and NZH conditions.

Prediction of variables of interaction-experience components									
Target variable	Predictor variable	BBC *n* = 43			NZH *n* = 42			BBC - NZH *df* = 83	
		β	*t*[a]	*SE*	β	*t*	*SE*	∆β	|*t*|
PQ	PD	-.55	***6.32	.09	-.65	***10.19	.06	.11	1.01
R^2_{BBC} = .57	POS	.46	***4.28	.11	.23	**2.78	.08	.23	1.70
R^2_{NZH} = .62									
UC	AI	.49	***4.10	.12	.38	*2.09	.18	.12	0.55
R^2_{BBC} = .57	POS	.40	***3.58	.11	.37	*2.29	.16	.03	0.17
R^2_{NZH} = .41									
PE	HQ	.24	1.72	.14	.62	***6.42	.10	-.38	*-2.25
R^2_{BBC} = .52	POS	.57	***4.43	.13	.34	***3.63	.09	.23	1.44
R^2_{NZH} = .83									
HQ	AES	.81	***9.57	.08	.68	***7.41	.09	.13	1.06
R^2_{BBC} = .74	POS	.09	0.80	.11	.29	**3.18	.09	-.20	-1.47
R^2_{NZH} = .83									
POS	AES	.58	***6.66	.09	.74	***13.32	.06	-.16	-1.53
R^2_{BBC} = .34									
R^2_{NZH} = .55									
Prediction of variables of interaction outcomes									
Target	Predictor	BBC			NZH			BBC - NZH	
		β	*t*	*SE*	β	*t*	*SE*	∆β	|*t*|
GOOD	HQ	.49	**3.18	.15	.73	***5.03	.14	-.24	-1.14
R^2_{BBC} = .45	PQ	.27	*1.97	.14	.25	1.65	.15	.02	0.08
R^2_{NZH} = .83									
BI	PE	.11	0.80	.14	.53	***3.99	.13	-.42	*-2.19
R^2_{BBC} = .37	TRU	.09	0.54	.16	.23	1.56	.15	-.15	-0.68
R^2_{NZH} = .44	UC	.50	**2.84	.18	.20	1.41	.14	.30	1.36
SAT	PE	.28	**2.58	.11	.42	*2.34	.18	-.15	-0.70
R^2_{BBC} = .67	PQ	.35	***3.35	.11	.28	1.53	.18	.08	0.38
R^2_{NZH} = .77	TRU	.32	**3.10	.10	.16	1.48	.11	.16	1.10
	UC	.15	1.47	.10	.36	**2.88	.13	-.21	-1.29
BEAU	HQ	.36	**3.18	.11	.52	***4.02	.13	-.16	-0.96
R^2_{BBC} = .45	POS	.24	1.67	.14	.27	*2.34	.11	-.03	-0.16
R^2_{NZH} = .74	PQ	.21	1.56	.13	.16	1.76	.09	.05	0.32

Note. AES: aesthetics. AI: adequacy of information. BEAU: beauty. BI: behavioural intention. GOOD: goodness. HQ: hedonic quality. PD: perceived disorientation. PE: perceived enjoyment. POS: positive affect. PQ: pragmatic quality. SAT: satisfaction. TRU: trust. UC: usefulness of content.

[a]Bootstrap, *N* = 5000.

*p < .05. **p < .01. ***p < .001.

With regards to the prediction interaction-experience components in the experimental conditions, the path from *hedonic quality* to *perceived enjoyment*, and the path from *positive affect* to *hedonic quality* lost significance in the BBC condition, whereas both paths remained significant in the NZH condition. The path coefficient between *perceived enjoyment* and *hedonic quality* was significantly different in the experimental conditions; *hedonic quality* was a high-impact and highly significant predictor of *perceived enjoyment* in the NZH condition, but its impact was smaller and non-significant in the BBC condition. Otherwise, the prediction of components of interaction experience was robust across the conditions. Note that several *t*-values approached significance, which suggests that a larger sample size could lead to some of the differences in path coefficients reaching significance. For example, *positive affect* was a significant predictor of *pragmatic quality* in both conditions. This relationship was stronger in the BBC condition and the difference between the conditions approached significance ($\Delta\beta = .23$, $t = 1.70$, $p = .09$).

With regards to the prediction of variables of interaction outcomes, *pragmatic quality* lost its significance in the prediction of *goodness* in the NZH condition, presumably due to a smaller sample size as a result of the separate analysis of groups. *Hedonic quality* was a high-impact and robust predictor of *goodness* in both conditions. *Perceived enjoyment* was a significantly stronger predictor of *behavioural intention* in the NZH condition. On the other hand, *usefulness of content* had a markedly higher path coefficient in the prediction of *behavioural intention* for the BBC group, but the difference did not reach significance. *Trust* lost its significance as a predictor of *behavioural intention* in both conditions, which may be attributed to reduced sample size.

Pragmatic quality lost its significance as a predictor of *satisfaction* in the NZH condition; however, the relatively high standard error associated with the path in the NZH condition and the lack of substantial difference from the highly significant regression coefficient in the BBC condition suggest that the drop in significance can be attributed to high standard error and drop in sample size. The impact of *trust* on *satisfaction* in the NZH group was half of that in the BBC group and lost significance, but the difference between the conditions was non-significant. In turn, *usefulness of content* lost significance in the prediction of *satisfaction* in the BBC condition, and its impact was less than half than in the NZH condition. Finally, *pragmatic quality* lost its significance as a predictor of *beauty* in both conditions, along with *positive affect* in the BBC condition. Model parameters including indicators of effect size were calculated after the removal of non-significant paths. Significant predictors of components of interaction experience are presented separately for the BBC and NZH conditions in Table 6.12.

Hedonic quality was not significantly predictive to *perceived enjoyment* in the BBC condition ($\beta = .24$, $t = 1.72$, $p > .05$, $f^2 = .08$), presumably due to the drop in sample sizes because of the separate analyses. Additionally, *positive affect* was not significantly predictive to *hedonic quality* in the BBC condition ($\beta = .09$, $t = 0.80$, $p > .05$, $f^2 = .02$). *Positive affect* and *hedonic quality* were positively related in the BBC condition too ($r = .52$, $p < .01$), but the path coefficient from *positive affect* to *hedonic quality* was reduced, because of the high amount of variance *perceived aesthetics* shares with both variables. The prediction of interaction-experience components was otherwise robust across the conditions. Most notably, *perceived disorientation* was a significant and negative predictor of *pragmatic quality*, *adequacy of information* was a

significant and positive predictor of *usefulness of content*, and *perceived aesthetics* was a significant and positive predictor of *hedonic quality* in both conditions.

Table 6.12. Significant predictors of interaction-experience components in both experimental conditions.

BBC					
Target variable (number of predictors)	Predictor variable	β	t^a	sr^2	Unique effect size (f^2)[b]
Pragmatic quality (2)	Perceived disorientation	-.55	***6.32	.29	.65
R^2 = .56	Positive affect	.46	***4.28	.20	.46
Usefulness of content (2)	Adequacy of information	.49	***4.10	.15	.34
R^2 = .57	Positive affect	.40	***3.58	.12	.29
Perceived enjoyment (1)	Positive affect	.69	***8.88	R^2 = .48	.90
Hedonic quality (1)	Perceived aesthetics	.86	***28.21	R^2 = .73	2.72
Positive affect (1)	Perceived aesthetics	.58	***6.66	R^2 = .34	.52
NZH					
Target variable	Predictor variable	β	t^a	sr^2	Unique effect size (f^2)[b]
Pragmatic quality (2)	Perceived disorientation	-.65	***10.19	.33	.87
R^2 = .62	Positive affect	.23	**2.78	.04	.12
Usefulness of content (2)	Adequacy of information	.38	*2.09	.08	.13
R^2 = .41	Positive affect	.37	*2.29	.11	.19
Perceived enjoyment (2)	Hedonic quality	.62	***6.42	.14	.83
R^2 = .83	Positive affect	.34	***3.63	.04	.25
Hedonic quality (2)	Perceived aesthetics	.68	***7.41	.21	1.23
R^2 = .83	Positive affect	.29	**3.18	.04	0.23
Positive affect (1)	Perceived aesthetics	.74	***13.32	R^2 = .55	1.23

[a]Bootstrap, N = 5000.

[b]f^2: 0.02 - 0.14 small, 0.15 - 0.34 medium and 0.35 - large.

*p < .05. **p < .01. ***p < .001.

Although the drop in sample size resulting from separate analysis in the experimental conditions may render some paths with smaller effect sizes non-significant, other significant paths indicate effects that are robust across the conditions. Differences of path coefficients between the experimental conditions allow for the examination of the model across two artefacts at different stages of adoption. Model parameters of significant predictors of outcome variables are presented separately for the experimental conditions in Table 6.13.

Table 6.13. Significant predictors of measures of interaction outcomes in the BBC and NZH conditions.

BBC					
Target variable (number of predictors)	Predictor variable	β	t^a	sr^2	Unique effect size (f^2)[b]
Satisfaction (3)	Pragmatic quality	.41	***3.88	.10	.28
R^2 = .66	Trust	.37	***3.31	.22	.36
	Perceived enjoyment	.31	**2.77	.05	.16
Behavioural intention (1)	Usefulness of content	.59	***5.01	R^2 = .35	.54
Beauty (2)	Positive affect	.30	*2.33	.07	.12
R^2 = .42	Hedonic quality	.44	***4.73	.14	.25
Goodness (2)	Hedonic quality	.49	**3.18	.17	.31
R^2 = .48	Pragmatic quality	.27	*1.97	.05	.09
NZH					
Target variable (N of predictors)	Predictor variable	β	t^a	sr^2	Unique effect size (f^2)[b]
Satisfaction (3)	Pragmatic quality	.25	*2.25	.03	.19
R^2 = .83	Usefulness of content	.35	**2.61	.10	.57
	Hedonic quality	.53	***4.38	.14	.82
Behavioural intention (2)	Usefulness of content	.28	*1.99	.07	.11
R^2 = .39	Perceived enjoyment	.47	***3.83	.19	.31
Beauty (2)	Positive affect	.26	*2.12	.03	.10
R^2 = .72	Hedonic quality	.63	***5.59	.16	.58
Goodness (1)	Hedonic quality	.89	***30.54	R^2 = .79	3.83

[a]Bootstrap, N = 5000.

[b]f^2: 0.02 - 0.14 small, 0.15 - 0.34 medium and 0.35 - large.

*p < .05. **p < .01. ***p < .001.

In general, prediction accuracy of outcome variables in the NZH condition was higher, regardless of the number of significant predictors of any given variable. In the model based on the pooled sample, the significant predictors of *satisfaction* were *pragmatic quality*, *usefulness of content*, *trust* and *perceived enjoyment*. *Usefulness of content* was not a significant predictor of *satisfaction* in the BBC condition (β = .15, t = 1.47, *ns*, f^2 = .04), but it was significant and had a large effect size in the NZH condition. The effect of *trust* on *satisfaction*, on the other hand, was significant in the BBC condition and represented a large effect size, but it was not significant in the

NZH condition (β = .07, t = 0.84, *ns*, f^2 = .03)[35]. After the removal of the non-significant path between *trust* and *satisfaction* from the model in the NZH condition, the effect of *pragmatic quality* on *trust* regained significance. *Pragmatic quality* was the strongest predictor of *satisfaction* in the BBC condition with a medium effect size (although trust represented a large effect size), whereas it was the least strong predictor in the NZH condition with a medium effect size. Nevertheless, *pragmatic quality* was significantly predictive to *satisfaction* in both conditions. The paths between variables of components of interaction experience and variables of interaction outcomes for both conditions are presented in Figure 6.5.

In the analysis of the pooled sample, *perceived enjoyment* was a stronger predictor of *satisfaction* than *hedonic quality*, and the effect of *hedonic quality* was fully mediated through *perceived enjoyment*. However, the same analysis in the NZH sample revealed that *hedonic quality* had a stronger direct effect on *satisfaction* (β = .69, t = 7.50, p < .001, sr^2 = .38) than *perceived enjoyment* (β = .60, t = 5.35, p < .001, sr^2 = .31), and the simultaneous inclusion of both predictors in the path model rendered the direct effect of *perceived enjoyment* non-significant (β = .08, t = .53, p > .05). In order to explore the cause of the opposite pattern in the analysis using the full sample, the effect of *hedonic quality* on *satisfaction* was tested in the BBC sample too. Analysis revealed that in the BBC sample, even with the removal of *perceived enjoyment* from the model, the direct effect of *hedonic quality* on *satisfaction* was non-significant (β = .08, t = 0.56, p > .05, f^2 = .01). Therefore, it can be concluded that the full mediation of the effect of *hedonic quality* on *satisfaction*

[35] These parameters were recalculated after the exclusion of *perceived enjoyment* and the inclusion of *hedonic quality* in the NZH condition.

through *perceived enjoyment* in the pooled sample was an artefact of two separate direct effects in the two conditions. In the BBC condition, *perceived enjoyment* was a significant predictor of *satisfaction* with a medium effect size, whereas in the NZH condition, *hedonic quality* was a significant predictor of satisfaction with a large effect size, and the effect of *perceived enjoyment* on *satisfaction* was non-significant with *hedonic quality* present in the model.

Perceived enjoyment was not a significant predictor of *behavioural intention* in the BBC condition ($\beta = .11$, $t = 0.80$, $p > .05$) and the path represented a small effect size ($f^2 = .02$); therefore, it is unlikely that it would reach significance even with a large increase in sample size. Conversely, *perceived enjoyment* was a high-impact and significant predictor or *behavioural intention* in the NZH condition, with a larger path coefficient and effect size than that of *usefulness of content*. *Trust* lost its significance as a predictor of *behavioural intention* in both conditions; however, it had a greater effect size in the NZH condition ($f^2 = .09$) than in the BBC condition ($f^2 = .01$), and the loss of the significance of the path in the NZH condition ($t = 1.56$, $p = .12$) is presumably the result of the drop in sample size. Overall, *trust* was more strongly related to *satisfaction* in the BBC group, and the results tentatively suggest that it may be related to *behavioural intention* in the NZH group.

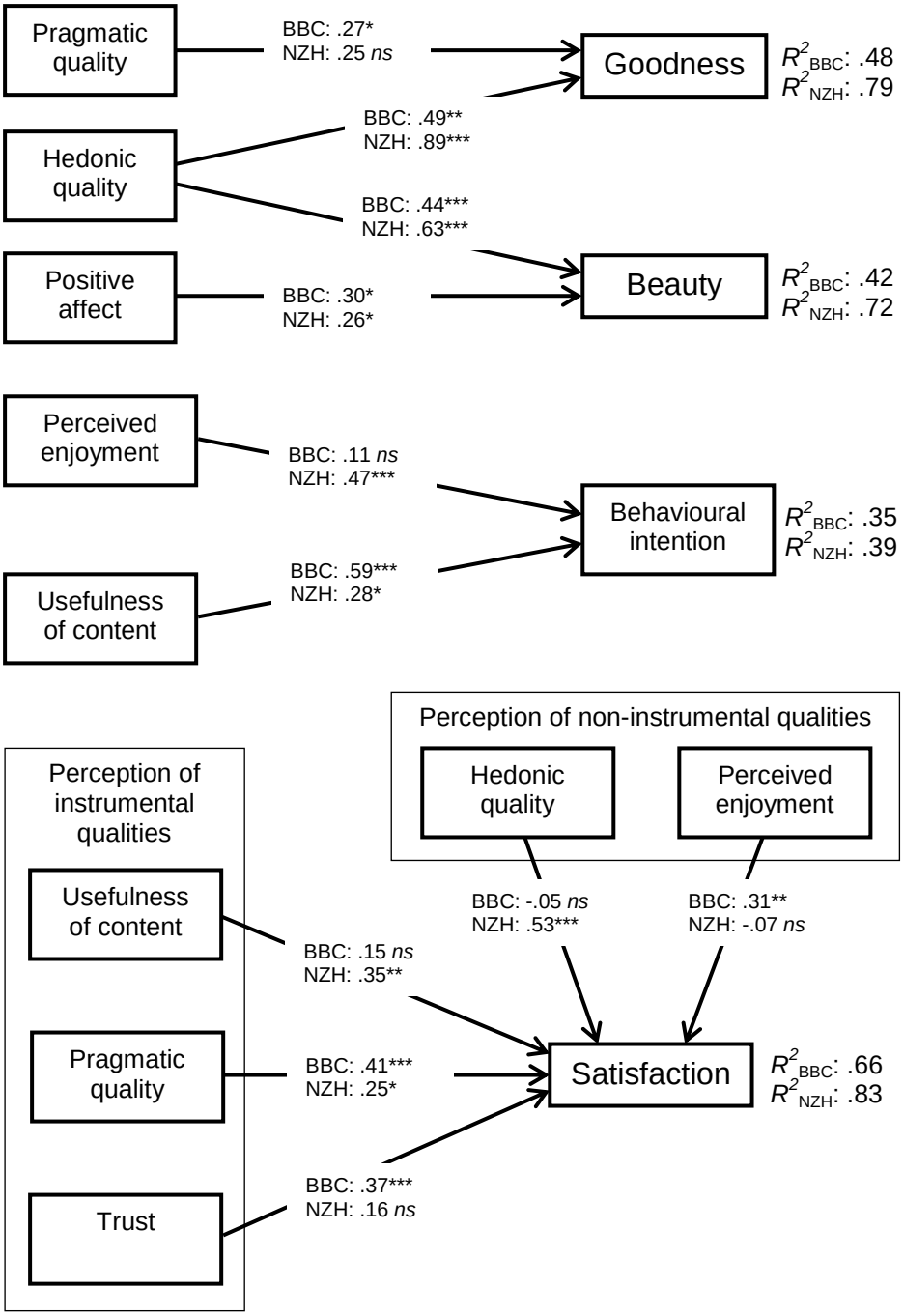


Figure 6.5. Prediction of variables of interaction outcomes in the BBC and NZH conditions.

Although *beauty* was predicted from *positive affect* and *hedonic quality* in both conditions, its prediction accuracy in the NZH condition exceeded that in the BBC condition by 30%, due to the larger path coefficient and effect size of *hedonic quality*. With all variables present in the model, the path from *pragmatic quality* to *beauty* became non-significant in the BBC condition too (β = .21, t = 1.56, *ns*, f^2 = .05). *Pragmatic quality* was not a significant predictor of *goodness* in the NZH condition (β = .25, t = 1.65, p = .10); however, it represented a medium effect size (f^2 = .21). An increase in sample size would likely to result in this relationship reaching significance. *Pragmatic quality* and *hedonic quality* were both significant predictors of *goodness* in the BBC condition. Overall, the results support the relationships between product attributes and product quality judgements as specified in Hassenzahl's (2003, 2004) model of interaction experience.

6.4.7. Tests of person-characteristic measures

6.4.7.1. Centrality of visual product aesthetics

Centrality of visual product aesthetics (CVPA) is defined as "the overall level of significance that visual aesthetics hold for a particular consumer in his/her relationships with products" (Bloch et al., 2003, p. 552). Research suggests that the aesthetic quality of products plays a more important role in purchase intention and attitudes toward a particular product for people characterised with high CVPA (Bloch et al., 2003). The measure of CVPA consists of three facets: value, acumen and response (for corresponding items, see Appendix 6.1). *Value* concerns the perceived personal and social importance of aesthetic design, *acumen* concerns the ability to recognise, categorise and evaluate the aesthetic design of products (i.e., 'aesthetic taste'), and *response* reflects the level of response to aesthetic design, in

other words, the importance of the visual design of products in general in driving purchase and use. Dimensions of CVPA were measured in the study to address individual differences in the salience of visual aesthetics among participants.

The examination of correlation coefficients between dimensions of CVPA and measures of aesthetic quality revealed that, although CVPA dimensions are significantly correlated (*r* values range from .32 to .51, $p < .01$), *value* and *response* were independent from ratings of *perceived aesthetics* and *beauty* at both times of measurement (*r* values range from .00 to .18, $p > .05$)[36]. Only the correlation between *acumen* and after-use *beauty* reached significance, but even in this case the two variables shared less than 5% variance ($r = .22$, $p = .045$). In summary, correlation analyses indicate that differences in CVPA among individuals do not lead to differences in ratings of aesthetics quality.

Perceived aesthetics was a direct, high-impact and significant predictor of *hedonic quality* and *positive affect* in the model. Although higher ratings of CVPA dimensions did not lead to higher ratings in aesthetic quality, it may be that the perceived aesthetic quality of user-interfaces is more salient in influencing affective reactions and the perception of products as being capable of promoting high-quality experiences for those who regard aesthetics more important in product design. Therefore, it is proposed that *CVPA value*, the perceived importance of visual design, and *CVPA response*, the level of response to aesthetic design, positively moderate the relationship between *perceived aesthetics* and *hedonic quality*; that is, higher levels of *value* and *response* are associated with a stronger connection between

[36] See Table 7.4 for each scale inter-correlation.

perceived aesthetics and *hedonic quality* (after use)[37]. Similarly, it is proposed that higher levels of *value* and *response* are associated with a stronger connection between *perceived aesthetics* and *positive affect*. Therefore, the following hypotheses are proposed:

Hypothesis 29: *CVPA value* is a positive moderator of the relationship between *perceived aesthetics* and *hedonic quality*.

Hypothesis 30: *CVPA value* is a positive moderator of the relationship between *perceived aesthetics* and *positive affect*.

Hypothesis 31: *CVPA response* is a positive moderator of the relationship between *perceived aesthetics* and *hedonic quality*.

Hypothesis 32: *CVPA response* is a positive moderator of the relationship between *perceived aesthetics* and *positive affect*.

In PLS analysis, the moderation effect of a metric moderator is expressed as products of the indicators of the predictor variable and the moderator variable (Henseler & Fassott, 2010). In order to avoid issues with multicollinearity in relation to higher-order product terms, the predictor variable (*perceived aesthetics*) and the moderator (*CVPA value*) were centred (Aiken & West, 1991). The results of the moderation effects of *CVPA value* on the prediction of *hedonic quality* and *positive affect* from *perceived aesthetics* are presented in Figure 6.6. Hypotheses 29 and 30 were not supported. However, the moderation effect of *CVPA value* on the path from *perceived aesthetics* to *hedonic quality* approached significance, $\beta = .27$, $t = 1.55$,

[37] After-use (T2) *hedonic quality* was used, because *perceived aesthetics* was measured only after use.

p = .06 (one-tailed). Similarly, the moderation effect of *CVPA value* on the path from *perceived aesthetics* to *positive affect* also approached significance, β = .29, *t* = 1.38, *p* = .09 (one-tailed). Hypotheses 31 and 32 were not supported. Because the moderation effects of *CVPA response* were very low and non-significant (*hedonic quality* as outcome variable: β = .00, *t* = 0.01, *p* = .50 [one-tailed]; *positive affect* as outcome variable: β = -.13, *t* = 0.59, *p* = .72 [one-tailed]), the moderation effects of *CVPA response* are not discussed any further.

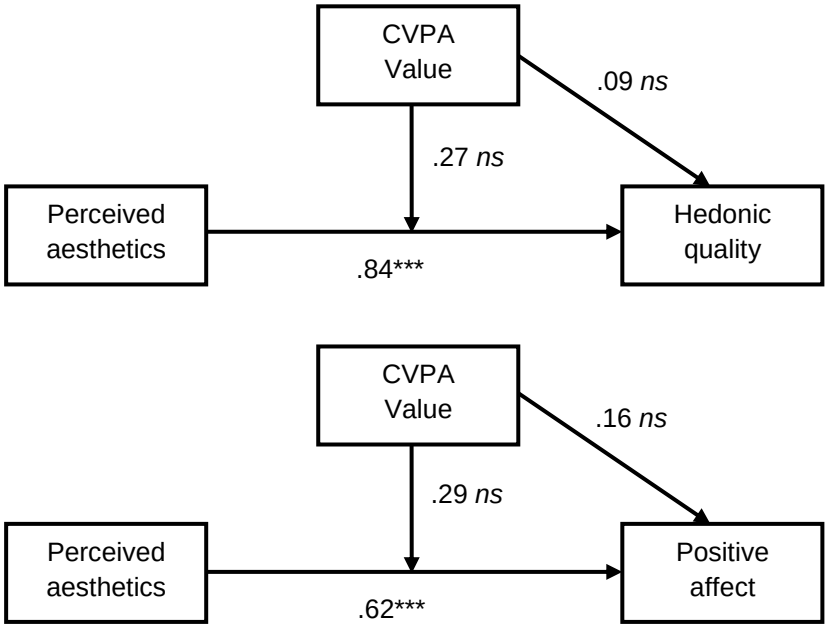


Figure 6.6. CVPA value as a moderator of the effect of perceived aesthetics.

Because the moderating effects of *CVPA value* did not reach significance, it needs to be concluded that the presence of moderating effects in the population is not supported. However, the findings tentatively suggest that the effect of perceived aesthetics on *hedonic quality* and on *positive affect* varies as a function of *CVPA value*. The high effect-size moderating effect of *CVPA value* suggests that a linear relationship between *perceived aesthetics* and *hedonic quality* was stronger in those

participants who attribute higher importance to design aesthetics in general (sr^2 = .06, f^2 = .36). The moderating effect of *CVPA value* in the analysis of the effect of *perceived aesthetics* on *positive affect* was of similar magnitude and represented medium effect size (sr^2 = .08, f^2 = .18), indicating that *perceived aesthetics* was more strongly related to the level of *positive affect* reported by participants who attribute higher importance to design aesthetics in general. Nevertheless, the lack of significance of parameter estimates does not allow the generalizability of the findings of the moderation analyses.

As Aguinis (2004) points out, moderation analysis in regression (and consequently in PLS) may suffer from low statistical power, which increases Type II error in the detection of moderation effects. Statistical power, in general, is sensitive to sample size; and the smaller the effect size, the more likely the particular effect will go undetected (Cohen, 1988). However, certain factors affect the power of moderation analysis in regression in particular (e.g., measurement error of latent variables and sample size; see Chin et al., 2003 and Aguinis, 2004). For example, Monte Carlo simulation research indicates that even medium moderation effect sizes may go undetected in samples as large as 120 from 65% to 9% of the cases (depending on the reliability of predictor variables) (Stone-Romero & Anderson, 1994). Monte Carlo simulations using PLS by Chin and colleagues (1996, 2003) demonstrated that in the case of approximately the same number of predictor-times-moderator indicator interaction terms[38] as present in the current study (six times six in the simulation, eight times four for *perceived aesthetics* and *CVPA value*), a sample size of *N* = 100

[38] The moderation effect is calculated as the product of each indicator (or item) of the predictor and moderator variables (Henseler & Fassott, 2010). For example, eight items for *perceived aesthetics* (predictor) and four items for *CVPA value* (moderator) results in 8*4 = 32 interaction terms.

is required for a β = .30 interaction effect reach significance of $p < .05$ (one-tailed). According to Aguinis (2004), sample sizes lower than 100 are inadequate for detecting moderation effects in general.

In the current study, moderation effect-sizes of *CVPA value* were medium to large. In a review of 261 moderation analyses involving categorical variables reported in the psychology literature, Aguinis and colleagues (2005) found a median effect size of $f^2 = .002$, which is considerably lower than the effect sizes in the current analysis. In a literature review of studies of moderation effects in information systems research, Chin et al. (1996) reported the average beta of moderating effects to be .10. The beta values of the moderation effects of *CVPA value* were markedly larger in the current study (.27 and .29). In summary, the findings and support from the literature tentatively suggest that *CVPA value* may moderate the relationship between *perceived aesthetics* and *hedonic quality*, and the relationship between *perceived aesthetics* and *positive affect* in the population, but these effects were not found to be significant, presumably due to low sample size; therefore, further empirical support is needed.

6.4.7.2. Spatial ability

The mental rotation of pseudo-3-dimensional polygons as a measure of spatial ability (Vandenberg & Kuse, 1978) was included in the study to explore if *spatial ability* as a person-characteristic variable is related to the level of *perceived disorientation* reported by participants after using news sites. As presented in Section 6.4.1, participants in the NZH condition ($M = 2.49$, $SE = .19$) reported significantly higher levels of *perceived disorientation* than participants in the BBC condition ($M = 2.00$, $SE = .15$), $t(83) = -2.04$, $p < .05$, presumably due to differences in familiarity with the

sites and/or differences in the navigation system of the two Web portals (Ahuja & Webster, 2001). The correlation between *spatial ability* and *perceived disorientation* was $r = -.18$, $p = .09$, which suggests that the extent to which users feel disoriented while using news sites (without externally provided information goals) is independent from their spatial ability, as measured by mental rotation of pseudo-3-dimensional polygons. Furthermore, *spatial ability* was uncorrelated with the number of home-page revisits ($r = .00$, $p = .98$) and the number of page visits ($r = .01$, $p = .96$), and *perceived disorientation* was uncorrelated with the number of home-page visits ($r = -.02$, $p = .84$) and the number of page visits ($r = -.02$, $p = .88$).

Juvina & van Oostendorp (2006) found that in Web navigation tasks, the so-called 'laborious navigation style', which is associated with a high rate of page revisits and high number of pages visited, was negatively related to *spatial ability* ($r = -.40$). According to the authors' interpretation, people navigating Web sites compensate for low *spatial ability* with increased effort. However, with no set-tasks in the current study (i.e., no pre-defined pieces of information to be found through browsing), users of news sites were free to pursue their own information goals emerging in the process of browsing. As a consequence, the number of pages visited and the number of home-page revisits were unrelated to the participants' level of *spatial ability*. The behavioural metrics of number of page visits and page revisits indicate disorientation in goal-directed use (McDonald & Stevenson, 1998), but in the task-free browsing context of the current study, these metrics were unrelated to *perceived disorientation*. Juvina and van Oostendorp (2006) found disorientation to be strongly related to perceived usability. The current findings support the strong connection between *perceived disorientation* and user-perceived usability (*pragmatic quality*) in a task-free context.

6.5. Discussion

The model of interaction experience with news sites was tested in a laboratory setting with two experimental groups. Participants in the BBC condition browsed the BBC news site for ten minutes, while participants in the NZH condition browsed the New Zealand Herald news site for the same amount of time. All participants were experienced users of the BBC site, but no participant had used NZH before taking part in the experiment. Measures of person characteristics were collected before interacting with the sites. Measures of interaction experience were collected following a self-paced presentation of five screen-shots, and after ten minutes of interaction with the sites.

Between-group differences were detected in participants' responses to interaction-experience measures. After use, BBC received significantly higher ratings on the interaction-outcome measures of *behavioural intention*, *satisfaction* and *goodness*, while judgements of *beauty* did not differ significantly between the groups. Additionally, participants in the BBC group provided significantly higher ratings of *usefulness of content*, *perceived enjoyment*, *trust* and *hedonic quality* than participants in the NZH condition. Similar to results in Study 1, there was no significant difference in the ratings of user-perceived usability between experienced and novice users of the sites, presumably because all participants were familiar with the general interface-design of news sites, and browsing without set tasks did not pose difficulties in navigation.

Hassenzahl (2004) found that perceptions of *beauty* and *hedonic quality* are stable over time, while perceptions of *pragmatic quality* and *goodness* judgements of interactive artefacts are more likely to be affected by hands-on experience.

Hassenzahl's findings were supported in the analysis of the stability of perceptions of product attributes and overall quality judgements in the BBC group. However, comparison of after-presentation and after-interaction ratings in the NZH group revealed that while there were no significant differences in the levels of *pragmatic quality* and *goodness*, *beauty* was rated significantly higher after use and the difference represented a medium effect size. This finding suggests that hands-on experience can provide more additional information about the appearance of a news site than about its usability when users browse a site without set tasks.

Differences in the stability of judgements between the BBC and NZH groups may be attributable to differences in mode of use (Hassenzahl & Ullrich, 2007). Although participants in both groups received instructions that promote use of the site in action mode (i.e., for exploration), people presumably use news sites in mixed mode, where exploratory behaviour (checking the news) alternates with pursuing emerging information goals (following up topics of interest). *Usefulness of content*, measured with a scale whose items referred to the relevance, currency and uniqueness of information presented by the site (see Appendix 6.1), was rated significantly higher in the BBC condition (see Table 6.6). Because BBC contains more relevant information to its regular users in the United Kingdom than the NZH, which they never used before and which is based in New Zealand, it is likely that using the BBC generates more information goals for its regular users than using NZH. As a consequence, users of BBC may have been likely to operate more often in goal mode during the experiment than users of NZH. Previous research demonstrated that usability can be more predictive to overall judgements in goal mode, and aesthetics can be more predictive to overall judgements in action mode (Hartmann et al., 2008). The stability of *pragmatic-quality* ratings in the NZH condition may be a consequence of the lack

of the importance of usability for users of NZH in action mode (provided that the level of usability is not excessively low), and the lack of stability in *beauty* ratings may be attributable to the increased importance of aesthetics in action mode. Conversely, the stability of *beauty* and *hedonic-quality* ratings in the BBC group may indicate that use experience did not provide further information about these aspects of the site. It is assumed that users of BBC operated more often in goal mode during their use of the site, and the lack of stability in *pragmatic quality* and *goodness* judgements may be attributed to usability's higher relevance to the quality of experience in goal mode.

Using the pooled sample, the test of the model of interaction experience with news sites formulated in Study 2 confirmed all structural paths of the model. However, contrary to the Hassenzahl's (2003, 2004) user-experience model, *pragmatic quality* was found to be a significant predictor of *beauty*. General-to-specific inference analyses found support for evaluative consistency in the inference of *pragmatic quality* from *beauty* through *goodness* at both times of measurement in the NZH condition. In the BBC group, evaluative consistency was supported in the inference of *pragmatic quality* before use, but the direct path between *beauty* and *pragmatic quality* was significant after use, supporting probabilistic consistency. The support for probabilistic consistency in the inference of *pragmatic quality* from *beauty* contradicts the findings of Hassenzahl and Monk (2010) and van Schaik et al. (under review), who found no direct link between the two variables after controlling for *goodness*. The lack of evidence of a direct connection between *beauty* and *pragmatic quality* after controlling for *goodness* in the BBC condition before use, and the presence of a direct connection after use suggests that use experience can be important in shaping the connection between beauty and usability in certain cases. Because the level of adoption and effects arising from the difference between artefacts were not controlled

separately in the present study, further evidence from novice and experienced users of news sites is needed to clarify the connection between beauty and usability in the context of news-site use. Additionally, general-to-specific inference analyses revealed that *hedonic quality* and *pragmatic quality* were independent from each other in both conditions at both times of measurement, after controlling for the effects of *beauty* and *goodness*.

The variables of perceptions of instrumental qualities were complemented with *trust* in the current study in an attempt to increase the prediction quality of *behavioural intention*, and to test the relationship between *trust* and *perceived usefulness*, proposed by Chen and Corkindale (2008) in their adaptation of the technology acceptance model to the application area of online news. Because no measurement method was available in the literature to address the perceived trustworthiness of an online news provider, a four-item *trust* scale was constructed, based on the dimensions of trust in news media identified by Kohring and Matthes (2007). The *trust* scale showed adequate psychometric properties: the latent variable explained 75% of variance in the items and all items had a loading above .80 on the latent variable, indicating good construct validity, internal consistency was high (*CR* = .92), cross-loadings were generally low, indicating adequate discriminant validity on the item level, and the square root of *AVE* exceeded all scale inter-correlations, indicating adequate discriminant validity on the construct level. *Trust* was confirmed as a significant predictor of *behavioural intention* in the context of news-site use; however, Chen and Corkindale's (2008) assumption that trust is an antecedent of perceived usefulness was not supported. Furthermore, *trust* was found to be a significant predictor with large effect size of *satisfaction* in the BBC group, but this path had a markedly lower impact and was non-significant in the NZH condition.

Satisfaction was included in the model among variables of interaction outcomes, because it is a widely-used concept in summarising users' general attitudes toward interactive products in e-commerce and marketing research, and among design practitioners. *Satisfaction* was the most accurately predicted variable of the measures of interaction outcomes, with substantial values of coefficients of determination[39] in both experimental conditions. *Pragmatic quality* was a significant predictor of *satisfaction* in both conditions; however, *trust* was significant in the BBC group, but not in the NZH group; and *usefulness of content* was only significant in the NZH condition. Presumably, participants in the NZH condition had no or very limited before-hand information about the trustworthiness of the site and were greatly reliant on inferring trustworthiness from other site characteristics, whereas all participants as regular users of BBC had information about its high reputation as a news provider. Separate analyses of the experimental conditions revealed significant differences in the prediction of *satisfaction* from the components of perception of non-instrumental qualities: in the NZH condition, *hedonic quality*, the perceived capability of a particular product to support high-quality experiences, was a high-impact and significant predictor of *satisfaction*, whereas in the BBC condition, *satisfaction* was directly predicted from *perceived enjoyment*, the extent to which the interaction with the site was pleasurable, although with a markedly lower path coefficient and effect size. Although *hedonic quality* and *perceived enjoyment* shared a large proportion of variance, *hedonic quality* had no significant effect on *satisfaction* in the BBC condition, even when *perceived enjoyment* was removed from the model. An experimental design with separate control of artefact properties and level of adoption

[39] According to Chin (1998), R^2 values for endogenous variables are characterised as: .67- substantial, .33 - moderate and .19 - weak.

would be needed to clarify the role of perceptions of non-instrumental qualities in the prediction of satisfaction with news sites.

The model of interaction experience with news sites that was tested in the current study is causative, which aims to account for the maximum amount of variance in variables of interaction outcomes. At the same time, the model aims to describe the relationships between constructs that are theoretically related to variables of interaction outcomes and are measuring aspects of interaction experience. Components of interaction experience, in turn, are predicted from measures of perceived product characteristics, which are variables that can be more closely interpreted in relation to product design than measures of components of interaction experience. For example, *perceived disorientation* can be conceptually connected to information architecture, the way the information is organised on a particular site, *adequacy of information* is related to the completeness, comprehensiveness and sufficiency of content that is being presented on the site, and *perceived aesthetics* is a gauge of the 'look and feel' of the site, which can be addressed in design.

Overall, the model based on the pooled sample accounted for a moderate to substantial proportion of variance in the outcome variables (coefficients of determination ranging from .43 to .74) and in the variables of components of interaction experience (coefficients of determination ranging from .42 to .79). Tests of hypotheses based on the findings of Study 2 confirmed the model, except for a significant path from *pragmatic quality* to *beauty*. Tests of difference between the model parameters of the experimental conditions revealed significant differences in the role of *perceived enjoyment* in predicting *behavioural intention*, and in the connection between *hedonic quality* and *perceived enjoyment*. The prediction of

components of interaction experience proved to be robust across different stages of adoption and between two different artefacts. In particular, the paths from *perceived disorientation* to *pragmatic quality*, from *adequacy of information* to *usefulness of content* (i.e., perceived usefulness), and from *perceived aesthetics* to *hedonic quality* were strong and significant in both conditions.

There were differences in the prediction of outcome variables between the experimental conditions, and several of these differences approached significance. For example, *trust* lost significance as a predictor of *behavioural intention* in the separate analyses of groups; however, it had a higher impact in the NZH condition, suggesting that perception of trustworthiness may be important in promoting use in new users (as it contributes to the satisfaction of experienced users). Of all the measures applied in the study, *trust* had the highest correlation with *adequacy of information* among the measures of perceived product characteristics in both subsamples (BBC: $r = .52$, $p < .001$; NZH: $r = .47$, $p < .01$), which suggests that providing comprehensive, complete and sufficient information may promote the perceived trustworthiness of news Web sites for experienced and novice users alike.

In general, using a larger sample in further studies would be useful for determining differences in model parameters between artefacts and different stages of adoption. The prediction accuracy of outcome variables was markedly higher in the NZH condition. In Study 2, the prediction accuracy of the outcome variables was higher in the case of adopted sites other than BBC, which suggests that the model can be applied more accurately to certain news sites or subtypes of sites (e.g., subscription-based versus advertisement-based business model).

CVPA and a mental rotation test were included to assess individual-difference variables in the model. Although the tests of moderation effects of *CVPA value* did not reach significance, presumably due to insufficient sample size required for the statistical analysis, the findings tentatively suggest that greater importance attributed to design aesthetics by individuals could be associated with a stronger link between the perceived aesthetic quality of a news site and (a) positive affect experienced during the interaction, and (b) ratings of the site's ability to promote high-quality experiences. *Spatial ability*, as measured by a mental rotation test of pseudo-3-dimensional polygons, was independent from *perceived disorientation* and navigation metrics involved in the experiment. However, the application of a wider range of behavioural metrics may reveal associations between disorientation and navigation style, and between spatial ability and navigation style in task-free browsing. Previous research demonstrated that spatial ability is strongly and positively associated with the successful completion of web-navigation tasks (Juvina & van Oostendorp, 2006; van Oostendorp & Juvina, 2007). Although the current study did not involve externally given tasks, this does not mean that participants did not pursue information goals. Future research may involve measurement of effectiveness and efficiency of completion of information goals defined by the participants themselves during their use of a news site (e.g., based on self-report). Then the relation between spatial ability and success in completion of navigation tasks might be demonstrated in a context without tasks that are externally given.

Generally, model testing would benefit from the use of multiple news sites, and recruiting a larger number of expert and novice users to each of the sites. The present study cannot distinguish between differences arising from level of adoption and from differences in artefact characteristics. The design of the current study,

however, was constrained by practical considerations. In particular, there were difficulties in the recruitment of users of a particular local news Web site in Study 1: despite advertising the study in two local newspapers and on the news site, and although monetary compensation was offered, very few participants volunteered for the study. Therefore, it was decided to use an artefact the users of which were easier to find (the BBC news site), but consequently it was extremely hard to recruit a sufficient sample of non-users. Finally, only users of the BBC site were recruited, and the New Zealand Herald was selected as a news site which is of similarly high-quality and written in English, but the participants were not likely to use it due to its apparent low relevance to a British audience. Nevertheless, the experiment demonstrated the model's ability to account for a sufficiently large amount of variance in endogenous variables related to interaction experience and interaction outcomes, the measures demonstrated sufficient sensitivity to distinguish between different conditions and the model was generally supported in a controlled setting, using both an adopted and a non-adopted artefact. Furthermore, the inclusion of CVPA demonstrated the feasibility and the potential usefulness of including measures of person characteristics as moderators in the model. With ample resources for recruitment, the model could be used to evaluate existing news sites with using expert and novice users as participants[40]. Additionally, an increase in sample size would allow for the testing of moderation effects of person characteristics (e.g., level of education, age and computer self-efficacy) on the strength of connections between the model's constructs, which can be used to define distinct groups of (potential) users.

[40] Without the measures of person characteristics, all the scales in the study after the use of a news site took approximately five minutes to complete for the participants.

Chapter 7

Conclusion and discussion

Contents

Chapter 7: conclusion and discussion

7.1. Overview

The aim of this chapter is to (a) summarise the three research stages presented in this thesis, (b) evaluate the model of interaction experience with news sites, discuss its limitations and implications to future work and (c) provide an example of how design guidance can be drawn from the model. The chapter starts with a brief summary of the research phases and discusses their relation to the main research questions. The main research questions identified in Chapter 1 were the following. (1) What are the factors of end-users' interaction experience with news sites? This research question was addressed in the review of models of technology acceptance and interaction experience in Chapter 2, and the exploratory study presented in Chapter 3. (2) How can the factors of interaction experience be measured? This research question was related to formulating a measurement model of interaction experience with news sites, which was presented in Chapter 4. (3) How are the different factors of interaction experience related? This research question was related to formulating a structural model of interaction experience with news sites, which was presented in Chapter 5. Study 3 was conducted to complement and test the model in a controlled experimental setting, and it was presented in Chapter 6. (4) How can the knowledge acquired by answering Research Questions 1, 2 and 3 be applied to guide system design? This research question is addressed in the current chapter. Following the summary of the research phases, the model of interaction experience with news sites is evaluated against the person-artefact-task framework and the interaction-experience models presented in Chapter 2. Limitations of the research work undertaken in the project and suggestions to future work are discussed. In order to address Research Question 4, the findings of the three studies are used to derive methodological and practical guidance for news-site

design. Finally, the contribution of the research project presented in this thesis to knowledge is summarised.

7.2. A summary of the research phases

7.2.1. Exploratory phase

The exploratory phase started with a literature review of models of technology acceptance and interaction experience. The aim of the literature review was to address Research Question 1 by identifying existing models and relevant research work to drive the development of the project. A person-artefact-task framework was adopted to guide the evaluation of the models. The technology acceptance model and its adaptations to the application area of the World Wide Web and to online news, and the unified theory of acceptance and use of technology were presented to identify determinants of intention to use and the various factors that were proposed and measured in previous research in relation to usability and technology acceptance. Models of interaction experience, such as the user-experience model and the model of components of user-experience, were presented to identify instrumental and non-instrumental aspects of interaction experience. The models of interaction experience were evaluated according to the person-artefact-task framework.

An exploratory study (Study 1) was conducted to complement the literature review in addressing Research Question 1 by identifying the aspects of interaction experience relevant to the use of news sites. An online questionnaire was used to collect information on Internet use and news-site use habits of users of Gazette Live, a local news site, and to recruit participants to a think-aloud study. The protocol analysis of verbal- and screen-capture recordings of users and non-users of Gazette Live

yielded five categories of experience: impression, content, layout, information architecture and diversion. Because participants had not been prompted as to which aspect of their experience to give account of during their use of Gazette Live, the collected categories were regarded as spontaneous, self-reported aspects of experience with a new site. The correspondence between the categories and aspects of interaction experience in previous research were discussed. Interaction-experience questionnaires collected after the think-aloud sessions revealed significant differences between users and non-users of Gazette Live, which indicated a need to account for the stage of adoption of news sites in later research phases. Exploratory correlation and regression analyses of the interaction-experience questionnaires found support for the user-experience model in the application domain of news sites.

7.2.2. Modelling phase

The aim of the modelling phase was to address Research Questions 2 and 3 by formulating a measurement model and a structural model of interaction experience with news sites. Existing measures were collected from previous research to address the measurement of aspects of interaction experience identified in Study 1. Additional variables were identified from the relevant literature and models to measure need fulfilment, affective reactions, and instrumental and non-instrumental aspects of interaction.

Study 2 applied an online questionnaire to collect responses to the selected measures of interaction experience from users of news sites. The reliability, validity and factor structure of the measures were assessed using factor analysis and partial-least-squares. Need-fulfilment scales were excluded from further analyses based on

psychometric considerations. To test the hypotheses related to the connections between the collected measures, a user-experience model, a technology acceptance model and an integrated experience-acceptance model were constructed. A model of interaction experience with news sites was constructed by including all variables present in the study and arranging them into three stages, according to the model of components of user-experience: variables of artefact characteristics, components of interaction experience and variables of interaction outcomes. Components of interaction experience were further divided into perception of instrumental qualities, perception of non-instrumental qualities, and affective reactions. Differences in the prediction of variables of interaction outcomes between users of BBC and other news sites were analysed to explore differences in the model attributable to differences between artefacts at the same stage of adoption.

7.2.3. Experimental phase

The primary aim of Study 3 was to test the model of interaction experience with news sites, constructed in Study 2, in a controlled experimental setting, using two artefacts at different levels of adoption, thereby addressing Research Question 3. Different levels of adoption were included to address the differences in the quality of interaction-experience between users and non-users of a news site in Study 1. The BBC site was used as the adopted news site, because it was used by the majority of participants in Study 2; therefore, it was the most realistic choice to recruit users of a particular news site. New-Zealand Herald was selected as the non-adopted site for the purpose of the experiment. The separate control of level of adoption and differences between artefacts need to be addressed in further research.

In line with the findings of Study 1, there were significant differences in responses to measures of interaction experience between the experimental groups, which indicated differences in experience across the experimental conditions and supported the sensitivity of several measures to the experimental manipulation. The stability of perceived product attributes and overall quality judgements between after-presentation and after-use of the news sites were investigated. Tests of hypotheses drawn from Study 2 were conducted to confirm the relationships between variables of artefact characteristics, components of interaction experience and variables of interaction outcomes. All hypotheses were confirmed, except for the hypothesis related to the relationship between user-perceived usability and overall beauty judgement. The relationship between user-perceived usability and overall judgement of beauty of the news sites was investigated by constructing general-to-specific inference models.

Measures of trust and satisfaction were added to the model and hypothesis tests were conducted to establish their relationships with interaction-experience and interaction-outcome variables. Differences in model parameters between the experimental conditions were tested to assess the stability of the model across different levels of adoption and different artefacts. The results supported the stability of the model to predict components of interaction experience from variables of artefact characteristics across the experimental conditions. Analysis revealed differences across the experimental conditions in the prediction of interaction-outcome variables from interaction-experience components, but overall the model was judged to be stable across conditions and it was able to account for a medium to substantial amount of variance in interaction-outcome measures. Finally, analyses of person-characteristic measures demonstrated the feasibility and potential usefulness

of including variables of individual differences in further research. In particular, the findings related to the moderation effects of CVPA, reported in Chapter 6, indicate that *CVPA value* may moderate the relationship between aesthetic- and hedonic-quality perceptions, and between aesthetic perceptions and affective reactions.

7.3. Evaluation of the model, limitations and future work

7.3.1. Evaluation framework

To address a comprehensive set of factors in relation to end-users' interaction with news sites, a person-artefact-task (PAT) framework was adopted for the purpose of the current thesis. Therefore, the model developed and tested in this thesis is evaluated using the PAT framework. The comprehensiveness of the model is assessed in comparison with the models of technology acceptance and interaction experience reviewed in Chapter 2. Note that the model built and tested in the current thesis is specific to the application area of news sites, whereas several models presented in Chapter 2 are broader in scope.

The PAT framework (based on Finneran & Zhang, 2003) specifies three distinct, but interacting, factors that affect the quality of experience: person (P), artefact (A) and task (T). Additionally, interaction-episodes with artefacts take place in a context (C). These components are treated as antecedents of interaction experience, which, in turn, leads to interaction consequences (see Chapter 2, Figure 2.2). Chapter 2 concluded that the most comprehensive model of interaction experience in terms of the level of correspondence with the PAT framework was the model of components of user-experience (Thüring & Mahlke, 2007). Therefore, following the hypothesis tests in Study 2, the variables in the study were arranged in the structure of a causative model, in line with the structure of the PAT framework and the model of

components of user-experience. The variables were arranged in three groups: interaction characteristics, components of interaction experience and interaction outcomes. Causative path models including all variables were created in Studies 2 and 3. In these models, variables of interaction outcomes were predicted from variables of interaction experience, which, in turn, were predicted from variables of interaction characteristics. Each group of variables is discussed separately in the following sections.

7.3.2. Interaction characteristics

Interaction characteristics in the present thesis theoretically involved artefact, person, task and context properties. Artefact properties related to interaction experience with news sites were identified in Study 1, and measures to these properties were identified and collected from existing research work in Chapter 4. Although these measures were not directly related to designable product characteristics, they were deemed as more directly interpretable in relation to system- and interface design than interaction-experience measures. Measures of artefact characteristics included *perceived aesthetics*, *perceived user-interface design*, *perceived disorientation*, *adequacy of information* and *accessibility*.

7.3.2.1. Artefact characteristics

Perceptions of instrumental and non-instrumental qualities, as components of interaction experience, were predicted from artefact characteristics and positive and negative affect in Study 2 (R^2 values ranged from .42 to .52, Median = .46). The prediction accuracy of measures of instrumental and non-instrumental qualities from *perceived aesthetics*, *perceived disorientation*, adequ*acy of information* and *positive affect* were higher in Study 3 (R^2 values ranged from .46 to .79, Median = .64).

Perceived disorientation was a significant predictor of *negative affect* (r^2 = .08) and *perceived aesthetics* was a significant predictor of *positive affect* (r^2 = .16) in Study 2. The prediction accuracy of *positive affect* from *perceived aesthetics* was markedly higher in Study 3 (r^2 = .42). Differences in predictive power between Study 2 and Study 3 were attributed to differences in the design of the two studies (online survey versus laboratory experiment). By comparison, in a laboratory experiment using two online stores as artefacts (Porat & Tractinsky, in press), perceived aesthetics[41] and usability together accounted for an average of 54% variance in the pleasure dimension of emotional responses (R^2 = .52 and .55 for the two artefacts, respectively).

The prediction of interaction-experience components from artefact characteristics is useful, because it allows for connecting accounts of experience and perceptions of artefact-qualities to properties of artefacts that can be manipulated in design. From a design perspective, the utility of interaction-experience models lies in their ability to describe and measure important aspects of interaction experience with interactive products, and describe which characteristics of particular products are important in driving experiences and to what extent.

There are many examples in the human-computer-interaction literature of systematic manipulation of artefact properties leading to systematic variation in interaction experience and interaction outcomes. For example, Thüring and Mahlke (2007) demonstrated that the manipulation of system properties of various interactive artefacts, such as the interfaces of mobile telephones and portable audio devices,

[41] The same perceived-aesthetics items were used in Studies 1 and 2.

may lead to differences in perceptions of instrumental and non-instrumental qualities, and may affect users' emotional responses. Van Schaik and Ling (2008) found that violation of presentation principles in screen design and information organisation of a Web site had negative effects on the quality of interaction, while adherence to these principles had a positive effect on ratings of pragmatic quality and goodness judgements. Similarly, van Schaik and Ling (2011) demonstrated that manipulations of the properties of a Web-based encyclopaedia related to usability may lead to differences in interaction-experience and technology-acceptance constructs, which, in turn, affect overall evaluations and technology acceptance.

The relationship between design elements, affective reactions and emotional responses has been investigated in previous research. For example, Kim et al. (2003) identified a set of emotional and aesthetic responses relevant to viewing homepages of Web sites, and assessed the relationship between emotions and design factors used by Web developers (e.g., shape, texture and colour of design elements) to derive design guidance. More recently, Cho et al. (2011) assessed the relationships between design features of mobile phones (e.g., shape, colour, size and ratio of design elements) and affective satisfaction in an online survey using participants from different age groups to guide product evaluation. However, these studies represent a fine level of granularity in terms of linking specific design-characteristics of interactive artefacts to emotional responses. Van Schaik and Ling (2011) separated the above two approaches to identifying higher-level concepts in the design of Web sites. The studies of Kim et al. (2003) and Cho et al. (2011) represent a screen-design-based approach, where factors of design are identified as physical design attributes of artefacts and the interaction between these attributes, which can be useful in deriving design guidelines for Web-developers. However,

Tractinsky et al. (2004) pointed out that the usefulness of design guidelines based on the screen-design-approach is limited, because of the large number of design options and their interaction with individual differences and Web-site type (e.g., retail, leisure and information-presenting) that need to be addressed. Instead, Hassenzahl (2004) and van Schaik and Ling (2011) argued for the necessity of theoretical models from an information-processing approach to guide the understanding of how interaction outcomes, such as overall judgements, satisfaction and use intentions, are formed. The studies presented in the current thesis were used to build and test a theoretical model of interaction experience for the application area of news sites, using an information-processing approach.

The model of interaction experience with news sites does not map the correspondence between design factors and affective reactions; instead, it connects factors of designable artefact-characteristics to a wider range of components of interaction experience and expresses the strength of these connections in terms of standardised regression coefficients, while the prediction accuracy of interaction-experience components are expressed as the proportion of variance explained in them by artefact characteristics. In the controlled setting applied in Study 3, the average variance accounted for in the variables of perceptions of instrumental and non-instrumental qualities by the three measures of artefact characteristics and *positive affect* in the pooled sample was 63% (*SD* = 13.89), which indicates a good overall fit of the model in accounting for the sources of interaction experience attributable to artefact characteristics. Furthermore, variables of perceptions of instrumental and non-instrumental characteristics were predicted from two variables each in Study 3, which indicates parsimony in the first stage of the model.

The comprehensiveness of the model of interaction experience with news sites in predicting interaction-experience components from artefact characteristics can be compared with the four interaction-experience models discussed in Chapter 2. The connection of components of interaction experience to artefact characteristics is explicit in the CUE model, and some of these connections had been empirically tested by Thüring and Mahlke (2007). The environmental-psychology model considers exclusively dimensions of affective reactions, which are predicted from perceived aesthetics and usability; therefore, the model of interaction experience with news sites is more comprehensive. However, the studies reported in the current thesis did not involve the systematic manipulation of artefact characteristics. Although Hassenzahl's user-experience model (2003), as presented in Chapter 2, Figure 2.7, includes product features such as content, presentation and functionality as antecedents of product character, the measurement of these features were not pursued in research work by the author applying the model. Instead, Hassenzahl and Tractinsky (2006) and Hassenzahl and Roto (2007) identified the fulfilment of universal psychological needs as a source of positive experiences with interactive technologies using the user-experience model. Based on these papers and the findings of Hassenzahl (2008) and Hassenzahl et al. (2010), need fulfilment was included in the measurement model of Study 2, but it was excluded from the structural model based on psychometric considerations. Finally, the model of user-interface quality assessment and the related experimental work (Hartmann et al., 2008) included variables that were identified as indicators of artefact characteristics and interaction-experience components in the current thesis (e.g., aesthetics, service- and information quality, engagement and usability), but these variables were not separated into the two groups above; rather they were treated as members of a

single set of determinants of overall preference and task performance. By separating determinants of interaction outcomes into the categories of interaction-experience components and artefact characteristics, the model of interaction experience with news sites accounted for a moderate to substantial amount of variance in interaction-experience components in a controlled laboratory setting applied in Study 3, thereby accounting for sources of interaction experience in general factors of artefact properties.

7.3.2.2. Person characteristics

Spatial ability and centrality of visual product aesthetics were included as measures of person characteristics in Study 3. Spatial ability, as measured by the mental rotation of pseudo-3-dimensional polygons, was found to be unrelated to user-perceived usability and to the participants' level of disorientation in the free-browsing context applied in Study 3. Although the tests of moderating effects of *CVPA value* were not statistically significant, the effect sizes associated with the moderating effects were moderate to large, and it was concluded that an increase in sample size would likely to result in the moderation effects reaching significance. The inclusion of person characteristics as moderators allows for the examination of interaction effects between person and artefact characteristics in the formulation of interaction experience.

As discussed in Chapter 2, theoretical extensions of the technology acceptance model and the unified theory of use and acceptance of technology incorporated variables of person characteristics as direct determinants of technology-acceptance constructs as moderator variables. These variables included, for example, intrinsic motivation, computer anxiety, age and level of experience. Previous research

demonstrated that employing measures of person characteristics can be useful in predicting, for example, user-perceived usability (e.g., Venkatesh, 2000). With regards to interaction-experience research, van Schaik and Ling (in press [a]) measured intrinsic motivation as a person-characteristic variable in an experiment addressing both experiential and cognitive variables in Web navigation, but its effect on experience (flow), task performance and task outcomes were not significant. However, the effect of intrinsic motivation on flow was significant in a follow-up study (van Schaik & Ling, in press [b]).

The model of user-interface quality assessment acknowledges users' background as influencing the judgement of product-quality attributes (see Chapter 2, Figure 2.10). Although Hartmann et al. (2007) found that participants with different backgrounds (technical and design students) produced different rankings of sites based on ratings of aesthetics, experimental work related to the model did not include the measurement of person characteristics. Participants of Hartmann et al. (2008) recommended expressively aesthetic (metaphor-based) versions of educational Web pages to be used by children, and classically aesthetic (menu-based) versions to be used by more mature target populations, which suggests that overall preference for Web sites with matching content but different interaction-design may be different in target populations. However, user-background of the actual participants, as a person-characteristic variable, was not measured in the experimental studies. Moreover, users' background and target audience applied in the studies were demographic details, rather than measures of personality, skills or motivation that may be relevant to users' experiences. The conceptualisation of and research related to Hassenzahl's user-experience model and the environmental-psychology model did not include variables of person characteristics to test their effect on

interaction-experience. The model of components of user-experience listed user characteristics as determinants of components of interaction experience, but the related research by Thüring and Mahlke (2007) did not include person characteristics. By contrast, Study 3 explicitly addressed and tested the person-characteristic variables of spatial ability and centrality of visual product aesthetics in the model of interaction experience with news sites, but the results were not statistically significant. Therefore, further research is needed to incorporate person characteristics into the model. Designers of news sites make efforts to accommodate different groups of users. For example, many news providers provide versions of their Web sites with different font sizes. Therefore, identifying relevant user characteristics and demonstrating their effect on interaction experience may provide useful information for designers. Examples of additional person-characteristic variables that could be addressed in future work include domain expertise and interest (Juvina & van Oostendorp, 2006), cognitive style (Gwizdka, 2008) and need for cognition (Amichai-Hamburger et al., 2007).

7.3.2.3. Task/context characteristics

With regards to task/context characteristics in models of interaction experience presented in the literature review, research using the user-experience model included the systematic manipulation of mode of use (Hassenzahl et al., 2002; Hassenzahl & Ullrich, 2007; van Schaik & Ling, 2009, 2011). Task characteristics and user goals are present in the conceptual representation of the model of user-interface quality assessment. However, 'scenario of use' (serious and leisure) was not experimentally manipulated or measured in the studies of Hartmann et al. (2008); participants merely provided preference judgements of versions of Web sites in hypothetical use-situations. An experiment reported in the same paper included the manipulation of

use context of mobile-telephone use: in the mobile condition, participants were walking in a room while interacting with a mobile telephone, while participants in the idle condition were stationary during the interaction. Although the conceptual model of components of user-experience acknowledges task/context characteristics, these characteristics were not addressed in research work. Finally, the environmental-psychology model does not address task/context factors in its conceptual framework or in research.

Task characteristics were not controlled in the studies presented in this thesis. Because the everyday use of news sites typically involves browsing for new information, participants in the three studies were not given specific information-retrieval tasks. However, people presumably use news sites to find specific pieces of information too. For example, people may look for results of sports events, specific political developments and for specific pieces of information regarding economic and cultural topics. A possible control for task characteristics could be achieved by manipulating mode of use. In each study of the present thesis participants were asked to interact with the news sites as they normally do; therefore, a context of action-mode was provided each time. However, it was proposed in Study 3 that regular users of BBC may have formed and pursued a larger number of information goals during the experiment than the novice users of NZH, because the content of BBC was more relevant to them as residents of the United Kingdom than the content of NZH; therefore, BBC potentially afforded the fulfilment of information needs to a greater extent. The inclusion of manipulation of context in further studies may shed light on how information goals affect the relationships between product characteristics and interaction-experience components, and between interaction-experience components and interaction outcomes. Another advantage of including

set tasks would be to examine the effects of interaction experience on task performance. For example, recent research (van Schaik & Ling, in press [a]) found that experience of use of a Web site (conceptualised as flow) had a significant positive effect of task performance, after controlling for the experimental manipulations of artefact and task complexity, which suggests that the quality of interaction experience may affect task performance directly. Furthermore, the results from previous research that included mode of use in relation to the user-experience model unanimously stress the importance of addressing context in interaction-experience research. However, the manipulation of mode of use was beyond the scope of the current thesis and needs to be addressed in future research.

Level of adoption as a context factor was included and investigated in two of the studies of the current thesis. Study 1 found differences in users' experience between novice and expert users of a news site. In particular, regular users of the site reported significantly higher levels of enjoyment, perceived the site as more aesthetic and of higher hedonic value, and judged the site to be better and more beautiful. Similarly, Study 3 found significant between-group differences in interaction experience measures across different levels of adoption. However, differences in experience stemming from the level of adoption and from artefact characteristics could not be separated in Study 3, due to the design of the experiment. Nevertheless, from the findings of Studies 1 and 3, it can be concluded that level of adoption plays an important role in users' interaction experience with news sites.

7.3.3. Components of interaction experience and interaction outcomes

Components of interaction experience in the model of interaction experience with news sites were divided into three groups: instrumental qualities, non-instrumental

qualities and affective reactions. Instrumental qualities included *pragmatic quality*, as a measure of user-perceived usability, and *usefulness of content*, as a measure of perceived usefulness. Additionally, the instrumental qualities variable group in Study 3 was complemented with the measurement of *trust*. Non-instrumental qualities included *hedonic quality* and *perceived enjoyment*. Affective reactions included *positive affect* and *negative affect* in Study 2. *Negative affect* was removed from the measurement model of Study 3, based on psychometric considerations.

The user-experience model distinguishes pragmatic and hedonic attributes, which were typically measured in research related to the model as users' perceptions of a particular product's pragmatic and hedonic quality, respectively. The role of affect (positive and negative) in the formulation of interaction experience was also addressed in research applying the user-experience model. Instrumental and non-instrumental components of interaction experience in research related to the components of user-experience model included usability and aesthetics, respectively. Measurement of emotional responses in research related to the components of user-experience model included physiological and behavioural correlates of emotions, and subjective assessment of affective reactions (valence and arousal) and emotional responses (cognitive-appraisal dimensions). Variables of instrumental and non-instrumental components of interaction experience are not separated into distinct groups in the model of user-interface quality assessment. Measures applied in related research include usability, aesthetics, users' perceptions of content- and service quality and, similar to trust in the current model, reputation/identity (Hartmann et al., 2008). The model of user-interface quality assessment does not include emotional responses or affective reactions of users of interactive artefacts. Finally, predictors of attitudes as the measure of interaction outcomes in the environmental-

psychology model include the affect and valence dimensions of users' emotional responses, perceived aesthetics and usability.

Outcome measures used in research related to the user-experience model included appeal and overall evaluative judgements of beauty and goodness. Intention to use was included among measures of interaction outcomes in relation to an integrated experience-acceptance model by van Schaik and Ling (2011). Research related to the components of user-experience model used overall ratings and preference ratings of interactive products. Studies related to the model of user-interface quality assessment used overall preference as a measure of interaction outcomes. The environmental-psychology model includes attitudes towards Web stores as a measure of approach/avoidance responses.

Measures of interaction outcomes in Study 2 included the overall quality judgements of *beauty* and *goodness*, generally used in research related to the user-experience model, and *behavioural intention*, which is widely applied in models of technology acceptance. *Satisfaction* as an additional outcome measure was included in Study 3. Therefore, the model of interaction experience with news sites contains overall evaluative judgements, intention to use and satisfaction as variables of interaction outcomes.

Prediction accuracy of outcome measures differed across the three studies presented in this thesis. Study 1 applied exploratory regression analyses on a small sample size (n = 25). Nevertheless, *pragmatic quality* and *hedonic quality* explained 84% of variance in *goodness* judgements, which corresponds to substantial amount of variance explained, according to Chin's (1998) guidelines for the assessment of the size of coefficients of determination. Furthermore, *hedonic quality* had a higher

impact on *goodness* (β = .76) than *pragmatic quality* (β = .28). These results were in agreement with those of Hassenzahl et al. (2002), who found that hedonic attributes were more strongly related to appeal ratings of Web sites in action mode than pragmatic attributes. *Hedonic quality* alone explained 50% of variance in *beauty* in Study 1. Overall, the exploratory regression analyses in Study 1 supported the user-experience model, and moderate to substantial amounts of variance were accounted for in the outcome measures of *beauty* and *goodness*.

Whereas in Study 1 participants used the Gazette Live news site in a controlled setting under think-aloud instructions, which incited them to concentrate solely on browsing the site for ten minutes and to constantly and consciously reflect on their experience of use, Study 2 applied an online questionnaire in which the participants were asked to browse a news site of their choice (for as long as they liked, but a recommended minimum duration of 'at least a couple of minutes' was provided in the instructions), then complete the questionnaires. Therefore, Study 2 involved more than one artefact and the setting of data collection was less controlled. As a consequence, variance explained in *beauty* and *goodness* were lower in Study 2 (40% and 47%, respectively), but still remained moderate in magnitude. As in Study 1, *hedonic quality* had a higher impact in predicting *goodness* (β = .49) than *pragmatic quality* (β = .28).

In Study 2, *positive affect* was a significant and positive predictor of *beauty* after controlling for *hedonic quality*, whereas *negative affect* was a significant and negative predictor of *goodness*, after controlling for *pragmatic quality*. These finding support the hygiene and motivator distinction (Zhang & von Dran, 2000) of interaction-experience components in relation to affect dimensions: experienced *negative affect*

(hygiene factor) may lead to decreased *goodness* evaluation, while experienced *positive affect* (motivator factor) may lead to increased *beauty* judgements.

Behavioural intention was also included among outcome measures in Study 2 and 34% of its variance was explained by *usefulness of content* and *perceived enjoyment*. Although *behavioural intention* was a significant predictor of frequency of use, *baseline use-frequency* had a higher impact on use frequency, presumably because the sites used in Study 2 were already adopted and regularly used by the participants. In agreement with the findings of van Schaik and Ling (2011), *hedonic quality* was a determinant of *perceived enjoyment*, but *hedonic quality* and *pragmatic quality* were not independent determinants of *behavioural intention*.

Data collection was more controlled in Study 3 than in Study 2. In Study 3, participants browsed one of the news sites (BBC or NZH) for ten minutes in a computer laboratory. *Goodness* and *beauty* judgements were elicited from participants both before interaction (after presentation of screenshots) and after interaction. Including the data from both experimental conditions, 70% of variance in *goodness* was explained after use. As in studies 1 and 2, *hedonic quality* had a higher impact in predicting *goodness* ($\beta = .65$) than *pragmatic quality* ($\beta = .26$), probably due to the context of action mode. Contrary to the user-experience model, *pragmatic quality* was also a significant predictor of *beauty* in Study 3 ($\beta = .18$); however, *hedonic quality* ($\beta = .46$) and *positive affect* ($\beta = .24$) had a stronger direct impact on *beauty*. *Trust* was included as an additional predictor of *behavioural intention* in Study 3. *Usefulness of content*, *perceived enjoyment* and *trust* accounted for 43% of variance in behavioural intention.

Satisfaction was added to the model as an additional outcome variable in Study 3. *Pragmatic quality*, *usefulness of content*, *trust* and *perceived enjoyment* accounted for 74% of variance in *satisfaction*. However, there were differences in the prediction of *satisfaction* between the experimental conditions. For example, for the non-adopted site (NZH) *hedonic quality*, the perceived capability of the site to support positive experiences was a significant predictor of *satisfaction*, whereas in the adopted condition (BBC), the effect of *perceived enjoyment* experienced during the use of the site was significant and the effect of *hedonic quality* was not. *Usefulness of content* was a significant predictor of *satisfaction* in the NZH condition, but not in the BBC condition. A possible explanation for this would be that the expected usefulness of content for the BBC site may have been high, as was the perceived level of *usefulness of content* (*M* = 5.98 on a 1-7 scale); therefore, us*efulness of content* acted as a hygiene factor for BBC and consequently did not contribute significantly to the participants' level of *satisfaction* (as the level of *usefulness of content* was perceived sufficiently high by participants). According to this argument, a lower level of perceived than expected *usefulness of content* would result in a negative relationship between *usefulness of content* and *satisfaction*. However, the expectations of participants regarding the quality of information presented by the news sites were not measured in Study 3. Finally, *trust* was a significant predictor of *satisfaction* in the BBC condition, but not in the NZH condition. Nevertheless, 83% of variance was explained in *satisfaction* in the NZH condition, and 66% in the BBC condition, which indicates a very good fit of the model in predicting users' satisfaction with news sites in both adopted and non-adopted contexts.

In summary, the model of interaction experience with news sites included each component of interaction experience present in the models reviewed in Chapter 2.

However, only the valence dimensions (positive and negative) of affective reactions were addressed in the current model. By contrast, research related to the components of user-experience model addressed a wide variety of measures of emotions, and research related to the environmental-psychology model also included the arousal and dominance dimensions of affect. On the other hand, restricting the measurement of emotions to the two affect dimensions with a psychometric instrument that was successfully applied previously in interaction-experience research yielded the benefit of parsimony in measurement; answers could be collected quickly and relatively effortlessly, and the findings could be related to previous research. As opposed to other models of interaction experience, the current model also included belief constructs of technology acceptance among instrumental and non-instrumental aspects of interaction experience. Furthermore, the studies related to the current model presented in this thesis involved a wider range of interaction-outcome measures than the rest of the models. Components of interaction experience accounted for moderate to substantial proportions of variance in measures of interaction outcomes in each of the three studies, which indicates a good overall fit of the model.

7.3.4. Applicability and comprehensiveness of the model

The applicability and comprehensiveness of the model of interaction experience with news sites is, in part, evaluated here against the PAT framework and research related to other models of interaction experience presented in detail in Chapter 2. A summary of studies related to models of interaction experience in Chapter 2 was presented in Appendix 2.1. Appendix 7.1 presents a summary of the three studies of the current thesis, based on the same aspects of evaluation: artefacts used in the studies, tasks carried out by the participants, measures of experience, person

measures, artefact measures, task/context measures, objective measures and measures of interaction outcomes. An overview of the summary table in Appendix 7.1 is presented in Table 7.1.

Table 7.1. A summary of types of variables applied in the studies of this thesis.

Source	Experience	Person	Artefact	Task/context	Objective	Outcomes
Study 1	✓		✓	**✓**		✓
Study 2	✓		✓			✓
Study 3	✓	✓	✓	**✓**	✓	✓

Note. Bold highlights indicate the experimental manipulation of a variable type.

With regards to artefacts, Study 1 used Gazette Live, a news Web site that predominantly provided local content. Participants in Study 2 were instructed to use a news site of their own choice. Although a wide variety of news sites were used, two thirds of the participants in Study 2 used the BBC news site before completing the online interaction-experience questionnaire. Nevertheless, the model was formulated based on data derived from a large sample of participants who reported their experiences with a self-selected variety of news sites. Finally, participants in Study 3 used the BBC and the NZH news sites, both of which provided predominantly national and international content. Overall, a wide variety of news sites was used in the studies presented in this thesis. Except for the NZH and a small number of sites selected by users in Study 2[42], all news sites involved in the studies were based in the United Kingdom, and all participants who reported their experience with these sites were recruited in the United Kingdom. Therefore, the model of interaction experience with news sites can be regarded as applicable to

[42] Fourteen participants out of 305 used a news sites that was not based in the United Kingdom in Study 2. One participant used a new site based in China, the remaining 13 used news sites published in English (for example, CNN).

news Web sites in the United Kingdom without any specific limitation. However, note that news sites were not randomly sampled in the studies.

With regards to participants, experienced users of the Gazette site in Study 1 were recruited from regular users of the site using advertisements in local newspapers and on the Gazette site, while novice users were recruited among staff and research students at Teesside University. In Study 2, participants were recruited among staff and students of various universities in the United Kingdom. Participants in Study 3 were recruited among staff and research students at Teesside University. Overall, the three studies included a large number of university students as participants, but the majority participants of Studies 1 and 3 were not students. Every participant in each study had to be a regular user of news sites. Therefore, the interpretations of the results of the studies are not restricted exclusively to the population of university students.

In each of the studies, participants were instructed to use the news sites as they normally would; therefore, an action-mode context of use was provided. In addition to browsing a news site, participants in Study 1 were under think-aloud instructions, which involved the verbalisation of their experiences. Although the veridicality of the recorded protocols may had been affected by the think-aloud task, the participants received verbal instructions, watched an example video and were provided practice to minimise the impact of reactivity in think-aloud protocols (see Chapter 3). In Study 2, participants were asked as part as the online questionnaire to interact with a news site; however, whether the participants actually used a news site before completing the interaction-experience questionnaire was not controlled (they were asked to provide the name of the site they used in a text box). In other words, Study 2 did not

control whether participants reported on a specific episode of interaction or they provided a cumulative account of previous experiences with a particular news site. In Study 3, participants were instructed to interact with a news site and explore its content for ten minutes, without any confounding tasks.

With regards to time of measurement, verbal and screen-capture data in Study 1 were collected during the participants' interaction with a news site, which was complemented with a set of interaction-experience measures after their interaction. Study 2 applied only measures that were collected after the participants' interaction with a news site. Study 3 applied measures of interaction experience both before interaction (i.e., after the self-paced presentation of five screenshots) and after interaction. The majority of the measures applied in the three studies were subjective (i.e., psychometric scales); however, the participants' browsing behaviour was recorded and analysed in Study 1 (e.g., time-line analysis of verbal protocols and location codes in the protocol transcripts) and Study 3 applied objective measures of browsing behaviour (number of pages visited and number of homepage revisits). As discussed in detail in the previous sections and presented in Table 7.1, characteristics of person, artefact and task/context variables were addressed in the three studies. Each study included measures of experience and interaction outcomes. The participants' level of adoption of a news site was controlled in Studies 1 and 3.

In comparison with studies related to the models of interaction experience presented in Chapter 2, Study 3 is unique in involving measurement and analysis of person, artefact and task/context characteristics, measures of experience, objective measures and measures of interaction outcomes within a single study (see Table 2.1

in Chapter 2 and Appendix 2.1 for an overview). Measures of person characteristics were scarcely used in the studies related to interaction-experience models reviewed in Chapter 2; only demographic details were analysed in two studies of the model of user-interface quality assessment (children or adults as target population and technical or design study as background). By comparison, the person characteristics collected in Study 3 were personality trait variables; centrality of visual product aesthetics addressed the participants' relationship to visual product-design, while spatial ability represented a more general trait.

In summary, the model of interaction experience with news sites and the three studies presented in this thesis agree with the PAT framework. Measures of users' experience were applied in each study. The three studies applied a wide range of measures of interaction experience, covering all aspects of experience that were addressed in previous research (e.g., usability, aesthetics, affect and need fulfilment). Data were collected before, during and after interaction. Both subjective (psychometric) and objective (navigation) measures were applied. In relation to the PAT framework, and in comparison to the reviewed models of interaction experience and the studies related to them, the current model and research work can be regarded as comprehensive. The main limitations are the lack of manipulation of artefact characteristics and mode of use. Additionally, further research is required to demonstrate the role of person characteristics in interaction experience with news sites. Increasingly comprehensive and accurate measurement of person-, artefact- and task/context characteristics in future research are expected to yield a more accurate account of interaction-experience components, which, in turn, demonstrated to be important predictors of interaction outcomes, such as intention to use, satisfaction and overall quality judgements. Implications on design from the model of

interaction experience with news sites and the findings of the three studies of this thesis are presented in the following sections.

7.4. Design guidance

The three studies presented in this thesis represent three different approaches to studying people's experiences with news sites. Although the main aim of the studies was to formulate and test a model of interaction experience with news sites, implications for system evaluation and design can also be drawn from them. In order to address Research Question 4, the following sections present a discussion of the applicability of the findings and the methods used in the studies to the evaluation and (re)design of news sites. Implications of methods used in the studies and the findings of the studies for design are presented separately.

7.4.1. Method implications

The analysis of the think-aloud protocols in Study 1 served as a basis for the selection of measures of interaction experience to be included in the measurement model in Study 2. Although news providers possess a wealth of practical knowledge that had been, and could be, used to identify aspects of experience with news sites (Chen & Corkindale, 2008), Study 1 focused on the users' interaction experience instead[43]. As opposed to, for example, focus-group settings in which the participants' post-hoc reports of experience and opinion are guided and biased by group facilitators, the concurrent think-aloud method applied in Study 1 collected self-reported aspects of experience that were unbiased by the researcher. Although the

[43] The purpose of the current thesis was to formulate and test a model of end-users' interaction experience with news sites rather than to test a model of experience based on the knowledge and intuition of news providers and developers of news sites.

think-aloud method is less controlled than surveys that apply standardised questionnaires on a large sample and experiments that involve manipulation of certain characteristic(s) of interaction, conducting think-aloud studies can provide valuable interaction-experience information with high detail and specific to a particular news site[44].

Study 2 applied an online survey method using standardised questionnaires. Although surveys are less controlled than laboratory settings and cannot provide the level of detail of think-aloud studies, they can provide large and representative datasets that can be used to draw predictive models to guide system (re)design. Although prediction-focused models drawn from such datasets may be theory-driven, as presented in Chapter 5, models can be flexible in terms of measurement specifications and goals. For example, measured aspects of experience and interaction characteristics can be tailored to a specific news site, based on theoretical considerations, such as a think-aloud study, and on managerial and design considerations (i.e., to test the performance of design aspects deemed important by a service provider). With regards to the flexibility of goals, aspects of experience and interaction characteristics measured on large samples can be used to predict dimensions of interaction (such as quality of experience or interaction outcomes) of specific interest. For example, a prediction target can be the satisfaction of users with a particular news site, another target can be the users' intention to revisit the site and yet another target may be the users' positive attitudes about the news provider. Measures can be included in surveys according to specific prediction goals. A utility

[44] Think-aloud studies have also been used to provide valuable usability information with high detail and specific to many different type of artefact.

of the model of interaction experience with news sites lies in providing an empirically tested basis for the selection of measures and their structural relationships. Guidance based on surveys can be expressed, for example, in the form of impact-performance matrices (Martensen & Grønholdt, 2003) to guide managerial and design decisions. Examples of impact-performance matrices drawn from the datasets of Studies 2 and 3 are presented in Section 7.4.3 of this chapter.

Study 3 applied a controlled laboratory setting to test the model of interaction experience with news sites using two artefacts at different levels of adoption. Experimental settings have the advantage of providing greater level of control than the think-aloud and online-survey settings, although they typically require more time and resources. Although the experiment in Study 3 was used to test the structural model formulated in Study 2 using news sites of two different news providers (cross-media testing), experimental settings can also be applied to aid news-site design. For example, experiments can be used to test different designs of the same online news medium simultaneously (within-medium testing) and to test different versions of the same medium before and after redesign (design-improvement testing). Furthermore, data collected in experimental settings can also be used in impact-performance analyses to guide further system (re)design. As demonstrated by the coefficients of determination acquired for the same measures applied in Studies 2 and 3, the controlled experimental setting of Study 3 yielded higher precision for predicting endogenous variables. On the other hand, survey settings have the advantage over experimental settings of facilitating the collection of larger datasets. Overall, experimental tests involve a higher cost and a lower number of participants than interaction-experience surveys, but provide greater control and precision for cross-media, within-medium and design-improvement testing.

In summary, the three methods applied in the studies presented in this thesis can be applied to evaluate and test online news sites. Each method has its advantages over the others in terms of control, precision, detail and representativeness, but applied together they can provide a wealth of empirical information to guide evaluation and design. The appropriate application of each method is well-documented in psychology textbooks and research literature. The current thesis provided an example of the application of each method to the subject area of interaction experience with news sites. The methods can be used to address the measurement and analysis of aspects of experience in a flexible manner through choice of focus in study-design (e.g., within-medium and design-improvement testing), the selection of measures for aspects of interaction to be involved (e.g., perceptions of artefact characteristics and affect) and the choice of focus in terms of prediction goals (e.g., satisfaction and continued use intention). Furthermore, think-aloud studies and evaluation of Web sites can be implemented in remote settings (Bastien, 2010) to include users who would not be available for laboratory experiments and participants representing a diverse population of users, for example, people with colour vision deficiencies (Power et al., 2009) or otherwise disabled users (Petrie et al., 2006). Although remote settings are less controlled (e.g., in terms of the environment and the computer hardware used by participants) and may involve technological problems in data collection (e.g., system crashes), they provide the benefit of involving a wide variety of technology configurations and environments, with the additional benefit of higher ecological validity than laboratory settings. The model of interaction experience with news sites formulated and presented in this thesis provides a theoretically grounded and empirically tested framework for studying end-users' interaction with news sites for research and design purposes. The next

section presents general design guidance and examples based on the three studies of this thesis.

7.4.2. Implications from the studies

Results of the analysis of think-aloud protocols in Study 1 supported the importance of content in driving the use and adoption of a news site (Chen & Corkindale, 2008); *content* was the largest and most prevalent category of units of thought. The units of thoughts related to the category *content* in Study 1 indicated the importance of relevance and timeliness of information presented on a news site. For example, a lack of diversity in news content (e.g., only crime stories reported), advertisements and outdated content elicited units of thought with negative valence, while stimulating (i.e., exciting and amusing) content and the presence of multimedia content usually elicited units with positive valence. It was confirmed in Studies 2 and 3 that the quality of information presented on news sites, conceptualised as *usefulness of content*, was a significant predictor of use intention. *Accessibility*, as a measure of service quality, and *adequacy of information*, as a measure of the completeness and comprehensiveness of content presented on the site, were significant predictors of *usefulness of content* in Study 2. *Trust* in Study 3, although significantly correlated with *adequacy of information* and *usefulness of content*, was an independent determinant of use intentions and *satisfaction*, which suggests that the perceived credibility and competence of a particular news provider (i.e., reputation) may not be exclusively derived from perceptions of information quality. For example, design aesthetics was a direct predictor of *trust* in an experiment using retail Web sites as artefacts while controlling for *usefulness*, *ease of use* and *customisation* (Li & Yeh, 2010). In summary, the quality of information as perceived by readers can be promoted by increased accessibility and adequacy of information. Information quality

is also positively associated with trust in the news provider as a determinant and it is an important driver of intention to use and satisfaction with news sites.

Because information architecture in the context of news-site use is related to issues with navigating complex, highly inter-connected structures of hyperlinks on frequently updated Web sites containing hundreds or thousands of pages, the users' perception of information architecture was measured with *perceived disorientation*. The readers' level of disorientation on news sites was strongly and negatively related to the perceived usability of the sites, which, in turn, was a predictor with medium effect size of *goodness* and *satisfaction*. Furthermore, disorientation was related to *negative affect*, which, in turn, had a direct negative effect on *goodness* evaluation of news sites. Therefore, it is important to address information architecture to improve the users' perception of a news site's usability. Good information architecture is expected to result in users experiencing less disorientation and less negative affect. The information architecture of news sites affects the efficiency of users finding the news they are looking for (Li, 2002). Moreover, according to Nielsen (2009), information architecture is still the most important factor causing usability issues for users of Web sites.

Information architecture can be improved by addressing the labelling system, organisation system and navigation system of Web sites (Rosenfeld & Morville, 2006). In particular, the labelling system can be improved by increasing the information scent of links and headings on a particular site, which refers to the extent to which readers can predict the content they will find behind hyperlinks (Nielsen, 2004). The probability of users feeling lost on a Web site can also be decreased by improving the organisation of information and by facilitating the ease of navigation

between pages presenting related content. Issues with information architecture can be identified with expert-based usability tests, such as 'cognitive jogthrough' (Rowley & Rhoades, 1992), and user-tests, such as card-sorting and rating (Bernard, 2000), but these methods can only be applied to small and medium sites (less than approximately 100 pages). News sites are typically large Web sites that may contain thousands of pages. However, cognitive computational modelling techniques, such as the Comprehension-based Linked model of Deliberate Search (CoLiDeS; Kitajima et al., 2000) have been proposed to offer an automated solution to analysing the information architecture of large Web sites.

A perceived measure of disorientation, such as the one applied in Studies 2 and 3, can be more accurate in assessing the level of users' disorientation while browsing news sites than behavioural measures, because the browsing behaviour of readers of news sites in action mode is influenced by factors other than disorientation, such as the exploration of site content (Smith, 1996). Research and models presented in Chapter 2 suggest that perceived usability and usefulness of news sites may be further improved by addressing various aspect of service quality. For example *accessibility* as a measure of service quality was included in Study 2 and proved to be a significant predictor of *usefulness of content*. Research related to models of technology acceptance and interaction experience proposed various service-quality measures (e.g., Lavie & Tractinsky, 2004; Yang et. al., 2005). These aspects related to usability and usefulness were not addressed in measurement in the current thesis, because they were not identified in the analysis of the think-aloud protocols in Study 1. However, including aspects of service quality in addition to *accessibility* in measurement may be useful to identify additional sources of usability and *perceived usefulness* of a news site to guide design.

Information architecture concerns the organisation and labelling of information elements, and navigation through the information structure of Web sites; however, it does not concern layout and graphics presentation. *Perceived user-interface design* was selected to address the measurement of the layout of the pages of news sites as perceived by users, and it was a predictor with medium effect size of *pragmatic quality*. Therefore, design changes in layout are expected to affect users' perception of the usability of the site and good layout is associated with higher levels of user-perceived usability. *Perceived aesthetics* and *perceived user-interface design* were highly correlated in Study 2, which suggests that layout is strongly related to the aesthetics of the interface. However, *perceived user-interface design* had a direct impact on *perceived enjoyment*, which, in turn, was an important predictor of people's intention to use a news site. The effect of *perceived aesthetics* on *perceived enjoyment*, on the other hand, was mediated through *positive affect* and *hedonic quality*. These findings suggest that although layout is related to aesthetics design, the distribution of content, navigation and graphic elements on the pages of a news site can directly influence the enjoyment experienced by readers during interaction. Therefore, addressing layout in design is important to promote both users' perceptions of usability and enjoyment of using news sites.

Detailed guidelines based on research for Web-page layout design are available and widely used in design practice (Koyani et al., 2004). Issues related to layout in Study 1 included, for example, page length, clustering of text, use of multimedia, consistency of presentation and page density (see Appendix 3.4). Conducting user-tests, such as think-aloud studies, can be useful to identify issues related to layout of particular news sites. Recording participants' eye movements while interacting with news Web sites using eye-tracking technology can be used to identify relevant

aspects of interface-design (Chu et al., 2009). Furthermore, think-aloud protocols can be combined with eye-tracking data to guide the evaluation of Web sites (Eger et al., 2007). With regards to aesthetics, practical guidance for design is far more limited (Kim et al., 2003); however, theoretical work and research on aesthetics in human-technology interaction (Lavie & Tractinsky, 2004) and emotional design (Norman, 2005) can provide useful implications to aesthetics design of news sites.

Each study in this thesis supports the importance of non-instrumental aspects of interaction with news sites. Both Study 1 and Study 3 found differences in users' experience between novice and experienced users of news sites. For example, novice participants in Study 1 were more likely to report issues with a news site (i.e., diversion) than expert users of the news site, who were more likely to report experiences with positive valence. The differences in the strength of path coefficients of the model in Study 3 imply that non-instrumental aspects of interaction, such as enjoyment and perceptions of hedonic quality, may play an even greater role in new users' experiences. The differences in path coefficients between users of BBC and users of other news sites in Study 2 imply that certain aspects of experience may be more relevant to users' judgement about particular news sites. The results of analysis of centrality of visual product aesthetics in Study 3 tentatively suggest that differences in experience can be expected due to characteristics of users. Furthermore, previous research discussed in this chapter emphasises the importance of task and context in the formulation of interaction experience with Web sites. It is therefore important to note or include person, artefact, task and context characteristics of a particular interaction episode when assessing quantitative data derived from surveys and experiments.

7.4.3. Impact-performance matrices

Design guidance based on quantitative interaction-experience data can be derived from impact-performance matrices (Martensen & Grønholdt, 2003), where the performance scores of variables predicting each outcome measure are presented in relation to their impact scores. An example of an impact-performance matrix based on users of BBC (*n* = 202) in Study 2 with *perceived enjoyment* as dependent variable is presented in Figure 7.1.

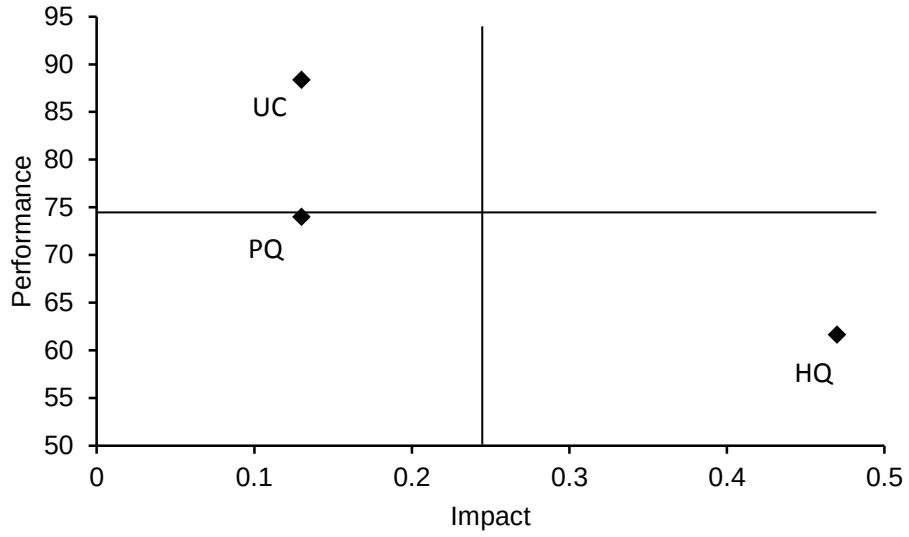


Figure 7.1. Impact-performance matrix of predictors of perceived enjoyment in the BBC group in Study 2.

Note. HQ: hedonic quality. PQ: pragmatic quality. UC: usefulness of content.

The X-axis represents the impact scores (standardised regression coefficients) of the predictors of the dependent variable. The vertical line represents the average impact of the three predictors (*M* = .24, *SD* = .20). The Y-axis represents the mean performance scores of participants on the independent variables, which are latent

variable scores acquired from running the PLS procedure[45]. For the purpose of the impact-performance analysis, the performance scores in Figure 7.1 have been transformed from a 1-7 to a 0-100 scale. The horizontal line represents the average performance of the independent variables. Additional parameters of the analysis are presented in Table 7.2.

Table 7.2. Parameters of the impact-performance analysis of the predictors of perceived enjoyment in Study 2.

Variable	Mean performance (*SD*)	Impact (β)
Hedonic quality (predictor)	61.61 (16.80)	.47
Pragmatic quality (predictor)	73.97 (17.52)	.13
Usefulness of content (predictor)	88.35 (16.05)	.13
Perceived enjoyment (criterion[a])	63.43 (19.44)	
Average impact of predictors (*SD*):	.24 (.20)	
Average performance of predictors (*SD*):	74.64 (13.38)	
Variance explained in perceived enjoyment:	36%	

[a]A criterion variable is a variable that is predicted from other variables.

Impact-performance matrices can be used to assess which areas of interaction-experience characteristics need to be addressed in design. The vertical (average impact) and horizontal (average performance) lines in impact-performance matrices divide the matrix into four cells. Each predictor variable can be placed in one of the quadrants. The lower-left quadrant contains variables that have a low performance, but have a low impact on the level of the dependent variable at the same time, relative to other predictors in the matrix. Low-performing significant determinants of interaction quality can be addressed in design. However, because the low-performing determinants in the bottom-left quadrant also have relatively low impact on interaction quality, it is not important to assign high priority to these aspects of

[45] Therefore, the impact scores of this analysis are exact linear combinations of the indicators of each construct, as opposed to the un-weighted averages of indicators.

interaction in system (re)design. In the present example, *pragmatic quality* had lower performance than the average; therefore, a recommendation can be made to improve *pragmatic quality*. Because *pragmatic quality* was predicted from *perceived disorientation* and *perceived user-interface design* in the model of interaction experience with news sites, it is proposed that addressing layout and information architecture in design may result in an increase in pragmatic-quality perceptions of users of BBC. However, it is not crucial to address *pragmatic quality* in design to enhance the level of enjoyment experienced by users of BBC while interacting with the site, because *pragmatic quality* was a relatively low-impact determinant of *perceived enjoyment* in the impact-performance analysis.

The upper-left quadrant in impact-performance matrices contains relatively high-performing, but low-impact variables. Variables that fall into this quadrant need not be addressed in design, because they are already performing relatively high and their impact is relatively low. It is suggested that high performance on variables in the upper-left quadrant, although not crucial, is beneficial for maintaining the high-quality of service. *Usefulness of content* in the present example is a relatively low-impact, although significant predictor of *perceived enjoyment*. The high performance of *usefulness of content* suggests that there is not much room for improvement, and its relatively low impact suggests that improvement would not result in a substantial change in the level of enjoyment experienced by users while interacting with the BBC news site. Therefore, no intervention in *usefulness of content* is advised to enhance *perceived enjoyment*. However, according to the model of interaction experience with news sites, users' perceptions of usefulness of content could be improved (or maintained) by addressing the completeness and comprehensiveness of information presented on the site (i.e., *adequacy of information*) and promoting accessibility.

The lower-right quadrant in impact-performance matrices contains relatively low-performing, but high-impact variables. Variables that fall into this area are the most important targets for design intervention. Variables in the lower-right quadrant are important predictors, but are not performing high relative to other predictors. In the current example, the relatively low performance of *hedonic quality* (*M* = 61.61, *SD* = 16.80) indicates that there would be room for improvement. According to the test results of the model of interaction experience with news sites, *hedonic quality* is strongly related to aesthetic quality. Therefore, addressing aesthetics in design is expected to yield changes in *hedonic-quality* perceptions. Because *hedonic quality* is a high-impact predictor of *perceived enjoyment*, improvement in *hedonic-quality* would be expected to result in increased enjoyment experienced by users of the BBC news site.

Finally, the upper-right quadrant of impact-performance matrices contains variables that are important predictors and perform high at the same time. Variables that fall into this quadrant represent strong areas of design that should be kept at a high level. No variable falls into this quadrant in the current example. The performance of *perceived enjoyment* as the dependent variable in the current impact-performance analysis (*M* = 63.43, *SD* = 19.44) indicates that there would be space for improvement. This analysis suggests addressing the determinants of *hedonic quality* in design as first priority. Addressing *pragmatic quality* is less important, but increasing users' pragmatic-quality perceptions of the site by addressing its determinants in design is expected to facilitate enjoyment. Note that the predictors only accounted for 36% of variance in *perceived enjoyment* in the current analysis, based on the data from Study 2. The next example is drawn from the data of Study 3, where the coefficients of determination of latent variables were higher.

Figure 7.2 presents an impact-performance analysis of the NZH group in Study 3 (*n* = 42) with *satisfaction* as the dependent variable, with additional results presented in Table 7.3. Note that *satisfaction* in the NZH group in Study 3 was predicted from the same variables as *perceived enjoyment* in the BBC group in Study 2. However, 83% of variance in *satisfaction* was explained by these variables (see Table 7.3 for the parameters). Furthermore, participants in the NZH group in Study 3 had not used the New Zealand Herald site before, as opposed to regular users of the BBC in Study 2.

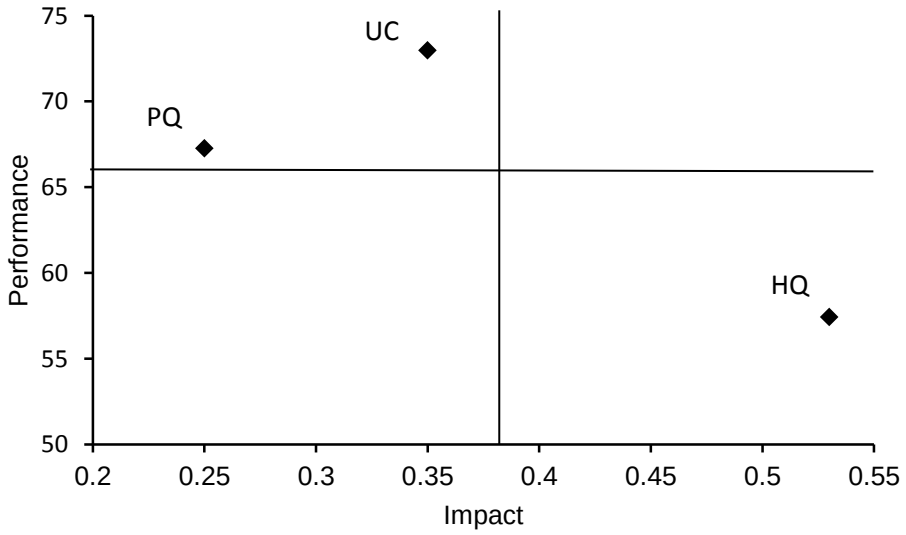


Figure 7.2. Impact-performance matrix of predictors of satisfaction in the NZH group in Study 3.

Note. HQ: hedonic quality. PQ: pragmatic quality. UC: usefulness of content.

Both *pragmatic quality* and *usefulness of content* were rated relatively high compared to *hedonic quality*; however, the performance scores indicate that there is space for improvement for both variables, especially *pragmatic quality*. Although the impact of *pragmatic quality* was low relative to the other predictors, it was still high compared to its impact in predicting *perceived enjoyment* in the Study 2 example presented in

Figure 7.1. *Hedonic quality* had the highest impact in predicting *satisfaction*, but it performed relatively low, which indicates room for improvement. The impact-performance analysis suggests that satisfaction of novice users of the NZH site could be improved predominantly by addressing the site's aesthetic appeal in design. Furthermore, according to Hassenzahl and Roto (2007), *hedonic-quality* perceptions play an important part in 'emotional product attachment', which suggests that people may be more likely to continuously use and describe themselves as users of a particular artefact (i.e., a BBC or an NZH reader) if they perceive the site to be highly conducive to having high-quality experiences.

Table 7.3. Parameters of the impact-performance analysis of the predictors of *satisfaction* in the NZH group in Study 3.

Variable	Mean performance (*SD*)	Impact (β)
Hedonic quality (predictor)	57.43 (21.66)	.53
Pragmatic quality (predictor)	67.27 (22.72)	.25
Usefulness of content (predictor)	72.98 (17.29)	.35
Satisfaction (criterion)	56.23 (24.52)	
Average impact of predictors (*SD*):	.38 (.14)	
Average performance of predictors (*SD*):	65.89 (7.87)	
Variance explained in satisfaction:	83%	

In summary, quantitative interaction-experience data can be used to construct impact-performance matrices, which can be used to identify areas of improvement for design. Impact-performance matrices are flexible in terms of prediction goals and can be drawn for any criterion variable with more than two significant predictors. It is important for the predictors in impact-performance matrices to be interpretable in terms of design. In the current examples, predictors were all interaction-characteristic variables that were connected to artefact characteristics. Although the impact-performance matrices presented here were drawn on the basis of the model

of interaction experience with news sites, any plausible set of predictors can be measured to predict criterion variables of particular interest. Constructing impact-performance matrices based on large datasets have the advantage that paths with low impacts are more likely to reach significance; therefore, more predictors can be included in the matrices. Large datasets are usually collected using survey methods, which are less likely to produce high values of coefficients of determination than controlled experimental settings. However, including larger sets of predictors in surveys is expected to result in better overall fit of prediction models in terms of the amount of variance accounted for by predictors in a particular criterion variable. Interaction-experience models, such as the model formulated and tested in this thesis, can be used to identify relevant sets of predictors of interaction outcomes and describe their relationships for the evaluation of (versions of) news sites.

7.5. Summary of contribution to knowledge

The main theoretical contribution to knowledge of the research project reported in this thesis is the model of interaction experience with news sites, presented in Figure 7.3. The structural model of interaction experience with news sites is causal; components of interaction experience are predicted from interaction characteristics; in turn, interaction outcomes are predicted from components of interaction experience. Interaction characteristics consist of user characteristics, perceived artefact characteristics and task/context characteristics. Level of adoption, as a task/context characteristic, influences perceptions of product characteristics and components of interaction experience (see Chapters 3 and 6 for between-group differences). Perceptions of artefact characteristics are direct determinants of interaction-experience components (see Chapters 5 and 6 for path analyses). User characteristics are proposed as potential moderators the effects of perceived artefact

characteristics on interaction-experience components (see Chapter 6 for moderation analyses); however, further research is needed to support this idea.

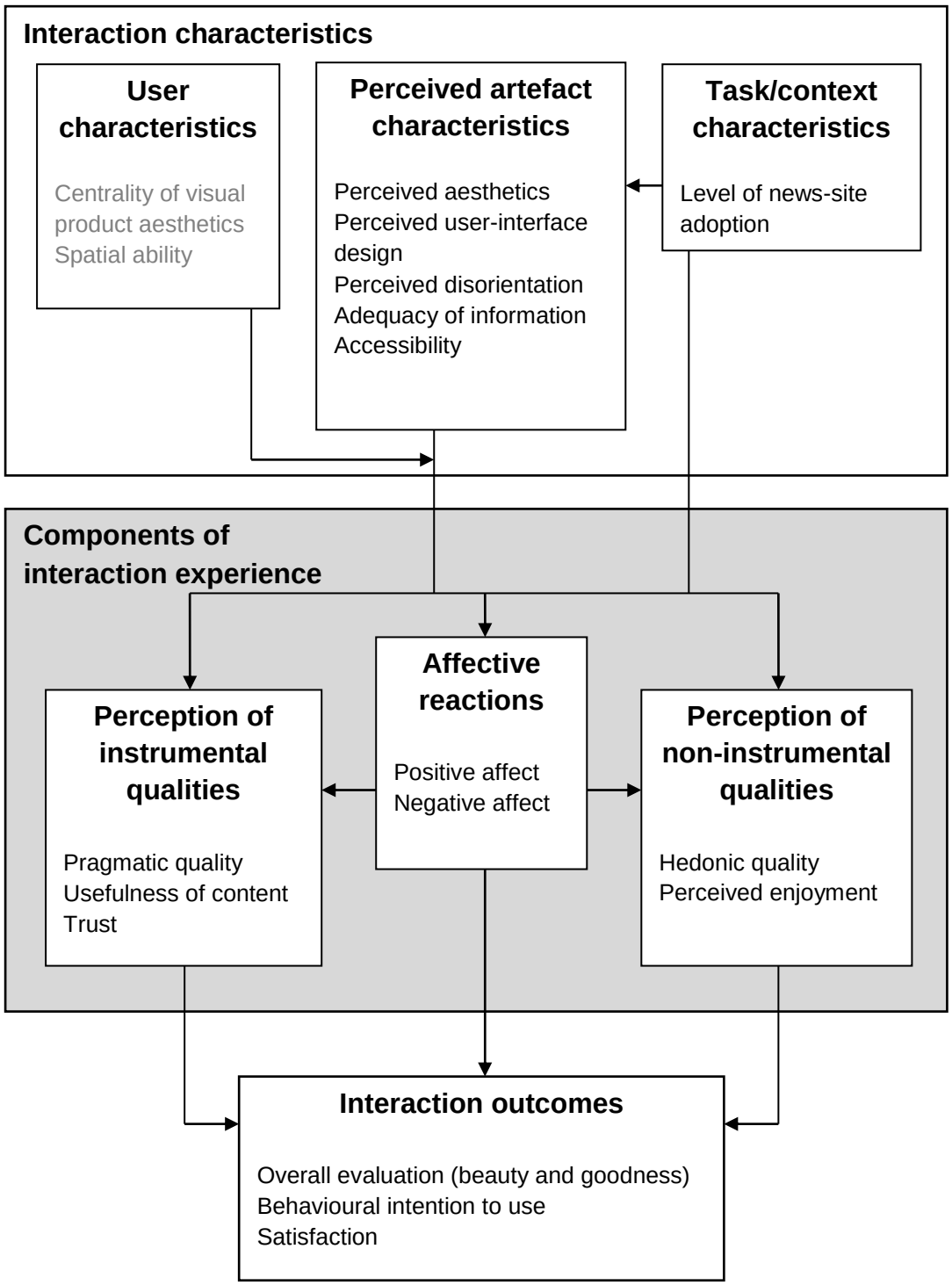


Figure 7.3. The model of interaction experience with news sites.

Components of interaction experience consist of perception of instrumental qualities, perception of non-instrumental qualities and affective reactions. Affective reactions influence perceptions of instrumental and non-instrumental characteristics, and interaction outcomes. Perceptions of instrumental and non-instrumental characteristics serve as direct determinant of interaction outcomes (see Chapters 5 and 6 for path analyses). Practical contributions of the research project, related to Research Question 4, are summarised in the current chapter (see Section 7.4). In addition, the measurement model that underlies the structural model constitutes a further contribution to knowledge. It specifies how the constructs in the structural model can be measured with good psychometric properties (see Chapter 4 for psychometric analysis).

Study 1 contributes to knowledge predominantly by identifying self-reported factors of interaction experience with news sites. Another important finding of Study 1 was the difference in interaction experience between regular users and non-users of a news site. Differences in interaction experience attributable to the difference in level of adoption were also supported in Study 3. Exploratory correlation and regression analyses in Study 1 demonstrated support for Hassenzahl's user-experience model.

Studies 2 and 3 contribute to knowledge predominantly by exploring and testing the structural relationships between measures of interaction experience, thereby formulating a model of interaction experience with news sites. The psychometric properties of measures of a wide range of factors (see Table 7.1) of interaction experience applied in various human-computer interaction models were assessed together in Studies 2 and 3 (see Chapters 4 and 6 for measurement models). The structural relationships between these measures were explored and tested in the

application area of online news in Studies 2 and 3 (see Chapters 5 and 6 for structural models). For example, in Study 2, *positive affect* was predominantly positively connected to perceptions non-instrumental characteristics and *beauty* evaluations of news sites, whereas *negative affect* was predominantly negatively connected to perceptions of instrumental qualities and *goodness* evaluations, which lends support to the distinction between hygiene factors and motivator factors as determinants of interaction experience (see Chapter 5). As another example, general-to-specific inference-perspective analyses in Study 3 supported the independence of *hedonic quality* and *pragmatic quality*, despite the high correlation between the two constructs. Notably, the inference-perspective model was not supported after use in the BBC condition.

A four-item measure of *trust* was introduced in Chapter 6. This scale was developed to measure trust in particular news providers, because existing measures of trust in news media only addressed trust in relation to the coverage of particular news events and news topics. The scale exhibited satisfactory psychometric properties and was confirmed as a direct determinant of *behavioural intention*, controlling for *usefulness of content* and *perceived enjoyment*, in the analysis of the pooled sample in Study 3. Furthermore, *trust* was a direct determinant of *satisfaction* with large effect size in the BBC group in Study 3.

Finally, Study 3 contributes to knowledge by introducing person-characteristic measures as moderators, in particular the *value* dimension of centrality of visual product aesthetics as a moderator between aesthetic perceptions and *hedonic quality*, and between aesthetic perceptions and *positive affect*. Although the moderating effects were not statistically significant, given the medium to substantial

effect sizes, the findings tentatively suggest the usefulness of including measures of person characteristics as moderators between perceived artefact characteristics and interaction-experience components in future research.

7.6. Final words

This chapter summarised the three research phases presented in this thesis and their findings, as well as the limitations of the studies and suggestions for future work. Implications for practical use of the research methods used in the three studies and the practical implications of findings of the studies were discussed. Next, examples of impact-performance analysis were presented to demonstrate how quantitative interaction-experience data can be used to identify areas of improvement and guide design decisions. Finally, the contribution to knowledge of the research project reported in this thesis was summarised.

The starting point of this project was to apply human-computer interaction and interaction-experience knowledge to the area of news-site use. It can be concluded that the main research questions of the thesis were addressed and answered by the three studies. The results of the studies presented in this thesis can be used to further develop the body of interaction-experience knowledge and provide guidance for design.

References

References

Aguinis, H. (2004). *Regression analysis for categorical moderators*. New York: The Guilford Press.

Aguinis, H., Beaty, J. C., Boik, R. J. & Pierce, C. A. (2005). Effect size and power in assessing moderating effects of categorical variables using multiple regression: a 30-year review. *Journal of Applied Psychology, 90*(1), 94-107.

Ahuja, J. S. & Webster, J. (2001). Perceived disorientation: an examination of a new measure to assess web design effectiveness. *Interacting with Computers, 14*(1)*,* 15-29.

Aiken, L. S. & West, S. G. (1991). *Multiple regression: testing and interpreting interactions*. Newbury Park: Sage.

Ajzen, I. (2002). Residual effects of past on later behaviour: habituation and reasoned action perspectives. *Personality and Social Psychology Review, 6*(2), 107-122.

Alben, L. (1996). Quality of experience: defining the criteria for effective interaction design. *Interactions, 3*(3), 11-15.

Allan, S. (2006). *Online news: journalism and the Internet*. Maidenhead: Open University Press.

Amichai-Hamburger, Y., Kaynar, O. & Fine, A. (2007). The effects of need for cognition on Internet use. *Computers in Human Behavior, 23*(1), 880–891.

Anderson, R. E. & Srinivasan, S. S. (2003). E-satisfaction and e-loyalty: a contingency framework. *Psychology & Marketing, 20*(2), 123-138.

Aranyi, G., Schaik, P. van & Barker, P. (2012). Using think-aloud and psychometrics to explore users' experience with a news Web site. *Interacting with Computers, 24*(2), 69-77.

References

Barker, P. G. & P. van Schaik (Eds.). (2010). *Electronic performance support: using digital technology to enhance human ability*. London: Gower Publishing.

Barwise, P., Elberse, A., & Hammond, K. (2002). Marketing and the Internet. In B. A. Weitz & R. Wensley (Eds.), *Handbook of Marketing* (pp. 527-557). London: Sage Publishing.

Bastien, J. M. C. (2010). Usability testing: a review of some methodological and technical aspects of the method. *International Journal of Medical Informatics, 79*(4), 18-23.

Bernard, M. (2000). Constructing user-centered websites: the early design phases of small to medium sites. *Usability News, 2*(1), Article 8, Retrieved December 8, 2011, from http://www.surl.org/usabilitynews/21/webdesign.asp

Bhattacherjee, A. (2001). Understanding information systems continuance: an expectation-confirmation model. *MIS Quarterly, 25*(3), 351-370.

Bloch, P. H., Brunel, F. F. & Arnold, T. (2003). Individual differences in the centrality of visual product aesthetics: concept and measurement. *Journal of Consumer Research, 29*(4), 551-565.

Boren, M. & Ramey, J. (2000). Thinking aloud: reconciling theory and practice. *IEEE Transactions on professional communication, 43*(3), 292-302.

Bruns, A. (2005). *Gatewatching: collaborative online news production*. New York: Peter Lang.

Card, S. K., Moran, T. P. & Newell, A. (1983). *The psychology of human-computer interaction*. London: Lawrence Erlbaum Associates.

Carroll, J. M. & Rosson, M. B. (1987). Paradox of the active user. In J. M. Carroll (Ed.), *Interfacing thought: cognitive aspects of human-computer interaction* (pp. 80-111). Cambridge, MA: MIT Press.

Carver, C. S. & Scheier, M. F., (1989). *On the self-regulation of behavior*. New York: Cambridge University Press.

Chen, Y. H. & Corkindale, D. (2008). Towards an understanding of the behavioral intention to use online news services: an exploratory study. *Internet Research, 18*(3), 286-312.

Chin, W. W. (1998). The partial least squares approach for structural equation modelling. In G. A. Marcoulides (Ed.), *Modern methods for business research* (pp. 295-336). Mahwah, NJ: Lawrence Erlbaum Associates.

Chin, W. W., (2000). *Frequently asked questions – partial least squares & PLS-graph.* Home page. Retrieved October 11, 2011, from http://disc-nt.cba.uh.edu/chin/plsfaq.htm.

Chin, W. W. (2010). How to write up and report PLS analyses. In V. E. Vinzi, W. Chin, J. Henseler & H. Wang (Eds.), *Handbook of partial least squares: concepts, methods and applications in marketing and related fields* (pp. 655-690). Berlin: Springer.

Chin, W. W. & Dibbern, J. (2010). An introduction to a permutation based procedure for multi-group PLS analysis: results of tests of differences on simulated data and a cross cultural analysis of the sourcing of information system services between Germany and the USA. In V. E. Vinzi, W. Chin, J. Henseler & H. Wang (Eds.), *Handbook of partial least squares: concepts, methods and applications in marketing and related fields* (pp. 171-193). Berlin: Springer.

Chin, W. W., Marcolin, B. L. & Newsted, P. R. (1996). A partial least squares latent variable modelling approach for measuring interaction effects. Results from a monte carlo simulation study and voice mail emotion/adoption study. In J. I. DeGross, S. Jarvenpaa & A. Srinivasan (Eds.), *Proceedings of the Seventeenth International Conference on Information Systems* (pp. 21-41). Cleveland, Ohio.

Chin, W. W., Marcolin, B. L. & Newsted, P. R. (2003). A partial least squares latent variable modelling approach for measuring interaction effects. Results from a Monte Carlo simulation study and an electronic mail emotion/adoption study. *Information Systems Research, 14*(2), 189-217.

Cho, V., Cheng, T. C. E. & Lai, W. M. J. (2009). The role of perceived user-interface design in continued usage intention of self-paced e-learning tools. *Computers & Education, 53*(2), 216-227.

Cho, Y., Park, J., Han, S. H. & Kang, S. (2011). Development of a web-based survey system for evaluating affective satisfaction. *International Journal of Industrial Ergonomics, 41*(3), 247-254.

Chu, S., Paul, N., & Ruel, L. (2009). Using eye tracking technology to examine the effectiveness of design elements on news websites. *Information Design Journal, 17*(1), 31-43.

Chuan-Chuan Lin, J. & Lu, H. (2000). Towards an understanding of behavioural intention to use a web site. *International Journal of Information Management, 20*(3), 197-208.

Clark-Carter, D. (2010). *Quantitative psychological research: the complete student's companion* (3rd ed.). Hove: Psychology Press.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37-46.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Csikszentmihalyi, M. (1990). *Flow: the psychology of optimal experience*. New York: Harpers Perennial.

Cyr, D., Head, M. & Ivanov, A. (2006). Design aesthetics leading to m-loyalty in mobile commerce. *Information & Management, 43*(8), 950-963.

Cyr, D., Hassanein, H., Head, M., & Ivanov, A. (2007). The role of social presence in establishing loyalty in eservice environments. *Interacting with Computers, 19*(1), 43-56.

Davis, F. D. (1986). *A technology acceptance model for empirically testing new end-user information systems: theory and results*. Doctoral dissertation, Sloan School of Management, Massachusetts Institute of Technology.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use and user acceptance of information technologies. *MIS Quarterly,13*(3), 319-340.

Davis, F.D., Bagozzi, R.P., & Warshaw, P.R. (1992). Extrinsic and intrinsic motivation to use computers in the workplace. *Journal of Applied Social Psychology, 22*(14), 1111–1132.

Davis, S. & Wiedenbeck, S. (2001). The mediating effects of intrinsic motivation, ease of use and usefulness perceptions on performance in first-time and subsequent computer users. *Interacting with Computers, 13*(5), 549-580.

De Angeli, A., Sutcliffe, A. G. & Hartmann, J. (2006). Interaction, usability and aesthetics: what influences users' preferences? *DIS 2006 Proceedings of the 6th conference on Designing Interactive systems*, 271-280.

Demir, E., Desmet, P. M. A. & Hekkert, P. (2009). Appraisal patterns of emotions in human-product interaction. *International Journal of Design, 3*(2), 41-51.

Desmet, P. & Hekkert, P. (2007). Framework of product experience. *International Journal of Design, 1*(1), 57-66.

Deuze, M., Bruns, A. & Neuberger, C. (2007) Preparing for an age of participatory journalism*. Journalism Practice 1*(3), 322-338.

Dickinger, A., Arami, M. & Meyer, D. (2008). The role of perceived enjoyment and social norm in the adoption of technology with network externalities. *European Journal of Information Systems, 17*(1), 4-11.

Dillon, A. (2001). Beyond usability: process, outcome, and affect in human computer interactions. *Canadian Journal of Information and Library Science, 26*(4), 67-69

Dishaw, M. T. & Strong, D. M. (1999). Extending the technology acceptance model with task-technology fit constructs. *Information & Management, 36*(1), 9-21.

Eger, N., Ball, L.J., Stevens, R. & Dodd, J. (2007) Cueing retrospective verbal reports in usability testing through eye-movement replay. *Proceedings of HCI 2007 (People and Computers XXI)*, 129-137.

Ericsson, K. A. & Simon, H. A. (1993). *Protocol analysis: verbal reports as data*. Cambridge, MA: MIT Press.

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: Sage.

Finneran, C. M. & Zhang, P. (2003). A person-artefact-task (PAT) model of flow antecedents in computer-mediated environments. *International Journal of Human-Computer Studies, 59*(4), 475–496.

Fishbein, M. & Ajzen, I. (1975). *Belief, attitude, intention and behaviour: an introduction to theory and research*. Reading, MA: Addison-Wesley.

Flavián, C. & Guerra, R. (2006). The choice of digital newspapers: influence of reader goals and user experience. *Internet Research, 16*(3), 231-247.

Flavián, C., Guinalíu, M. & Guerra, R. (2006). The role played by perceived usability, satisfaction and consumer trust on website loyalty. *Information & Management, 43*, 1-14.

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin, 76*(5), 378–382.

Ford, G. T. & Smith, R. A. (1987). Inferential beliefs in consumer evaluation: an assessment of alternative processing strategies. *Journal of Consumer Research, 14*(3), 363-371.

Fornell, C. & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: algebra and statistics. *Journal of Marketing Research, 18*(3)*,* 328-388.

Foust, J. C. (2009). *Online journalism: principles and practices of news for the Web* (2nd ed.). Scottsdale, Arizona: Holcomb Hathaway.

Freelon, D. G. (2010). ReCal: intercoder reliability calculation as a Web service. *International Journal of Internet Science, 5*(1), 20-33.

Gefen, D. (2002). Reflections on the dimensions of trust and trustworthiness among online consumers. *ACM SIGMIS Database, 33*(3), 38-53.

Gefen, D., Karahanna, E. & Straub, D. W. (2003). Trust and TAM in online shopping: an integrated model. *MIS Quarterly, 27*(1)*,* 51-90.

Gilhooly, K. & Green, C. (1996). Protocol analysis: theoretical background. In J. T. E. Richardson, (Ed.), *Handbook of qualitative research methods for psychology and the social sciences* (pp. 43-54). Leicester: British Psychological Society.

Gwizdka, J. (2008). Revisiting search task difficulty: behavioral and individual difference measures. *Proceedings of the 71st Annual Meeting of the American Society for Information Science and Technology (ASIST 2008)*. Silver Spring, Maryland: American Society for Information Science and Technology.

Ha, H-Y (2006). An integrative model of consumer satisfaction in the context of e-services. *International Journal of Consumer Studies, 30*(2), 137-140.

Haak, M. J. van den, Jong, M.D.T. de, & Schellens, P.J. (2004). Employing think-aloud protocols and constructive interaction to test the usability of online library catalogues: a methodological comparison. *Interacting with Computers, 16*(6), 1153-1170.

Hall, J. (2001). *Online journalism: a critical primer*. London: Pluto Press.

Hartmann, J., Sutcliffe, A. & De Angeli, A. (2007). Investigating attractiveness in web user interfaces. *CHI 2007 Proceedings of the SIGCHI conference on Human factors in computing systems*, 387-396.

Hartmann, J., Sutcliffe, A. & De Angeli, A. (2008). Towards a theory of user judgement of aesthetics and user interface quality. *ACM Transactions on Computer-Human Interaction, 15*(4), 1-30.

Hassenzahl, M. (2003). The thing and I: understanding the relationship between user and product. In M. Blythe, C. Overbeeke, A. Monk & P. Wright (Eds.), *Funology: from usability to enjoyment* (pp. 31-42). Dordrecht: Kluwer.

Hassenzahl, M. (2004). The interplay of beauty, goodness, and usability in interactive products. *Human-Computer Interaction, 19*(4), 319-349.

Hassenzahl, M. (2006). Hedonic, emotional and experiential perspectives on product quality. In C. Ghaoui (Ed.), *Encyclopedia of human-computer interaction* (pp. 266-272). Hershey PA: Idea Group Reference.

Hassenzahl, M. (2008). User experience (UX): towards an experiential perspective on product quality. *IHM 2008, Proceedings of the 20th International Conference of the Association Francophone d'Interaction Homme-Machine*, 339, 11-15.

Hassenzahl, M., Diefenbach, S. & Göritz, A. (2010). Needs, affect, interactive products - facets of user experience. *Interacting with Computers, 22*(5), 353-362.

Hassenzahl, M., Kekez, R. & Burmester, M. (2002). The importance of a software's pragmatic quality depends on usage modes. In H.Luczak, A. E. Cakir, & G. Cakir (Eds.), *Proceedings of the 6th international conference on Work With Display Units (WWDU 2002)* (pp. 275-276). Berlin: ERGONOMIC Institut für Arbeits- und Sozialforschung.

Hassenzahl, M. & Monk, A. (2010). The inference of perceived usability from beauty. *Human-Computer Interaction, 25*(3), 235-260.

Hassenzahl, M. & *Roto*, V. (*2007*). Being and doing: a perspective on user experience and its measurement. *Interfaces, 72*, 10-12.

Hassenzahl, M. & Tractinsky, N. (2006). User experience - a research agenda. *Behaviour and Information Technology, 25*(2), 91-97.

Hassenzahl, M. & Ullrich, D. (2007). To do or not to do: differences in user experience and retrospective judgments depending on the presence or absence of instrumental goals. *Interacting with Computers 19*(4), 429–437.

Hayes, A. F. & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures, 1*(1), 77-89.

Haynes, R. M. & Thies, E. A. (1991) Management of technology in service firms. *Journal of Operations Management, 10*(3), 388–397.

Heijden, H. van der (2004). User acceptance of hedonic information systems. *MIS Quarterly, 28*(4), 695-704.

Hekkert, P. (2006). Design aesthetics: principles of pleasure in product design. *Psychology Science, 48*(2), 157-172.

Henseler, J. & Fassott, G. (2010). Testing moderating effects in PLS path models. In V. E. Vinzi, W. Chin, J. Henseler & H. Wang (Eds.), *Handbook of partial*

least squares: concepts, methods and applications in marketing and related fields (pp. 655-690). Berlin: Springer.

Henseler, J., Ringe, C. & Sinkovics, R. (2009). The use of partial least squares path modelling in international marketing. In T. Cavusgil, R. Sinkovics & P. Ghauri (Eds.), *New challenges to international marketing*. *Advances in International Marketing, Vol. 20* (pp. 277-319). London: Emerald.

Hudlicka, E. (2003). To feel or not to feel: the role of affect in human-computer interaction. *International Journal of Human-Computer Studies, 59*(1-2), 1-32.

Internet World Stats (2011). *Internet and Facebook usage in Europe*. Retrieved December 20, 2011, from http://www.internetworldstats.com/stats4.htm

ISO (1998). ISO 9241: Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on usability. Geneve: International Organization for Standardization.

Jaspers, M. W. M., Steen, T., Bos, C. van den & Geenen, M. (2004). The think aloud method: a guide to user interface design. *International Journal of Medical Informatics, 73*(11-12), 781-795.

Juvina, I. & Oostendorp, H. van (2006). Individual differences and behavioural metrics involved in modelling web navigation. *Universal Access in Information Society, 4*, 258–269.

Kahneman, D. (1999). Objective happiness. In D. Kahneman, E. Diener & N. Schwarz (Eds.), *Well-being: the foundations of hedonic psychology* (pp. 3-25). New York: Russell Sage Foundation.

Kahneman, D. (2000). New challenges to the rationality assumption. In D. Kahneman & A. Tversky (Eds.), *Choices, values and frames* (pp. 758-774). New York: Russell Sage Foundation.

Karlsson, M. (2011). The immediacy of online news, the visibility of journalistic processes and a restructuring of journalistic authority. *Journalism, 12*(3), 279-295.

Kaynar, O. & Amichai-Hamburger, Y. (2008). The effects of need for cognition on Internet use revisited. *Computers in Human Behaviour, 24*(2), 361-371.

Keppel, G. (1982). *Design and analysis: a researcher's handbook* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Kim, J., Lee, J. & Choi, D. (2003). Designing emotionally evocative homepages: an empirical study of the quantitative relations between design factors and emotional dimensions. *International Journal of Human-Computer Studies, 59*(6), 899-940.

Kirakowski, J. & Corbett, M. (1993). SUMI: the software usability measurement inventory. *British Journal of Educational Technology, 24*,(3), 210–212.

Kitajima, M., Blackmon, M.H., & Polson, P.G. (2000). A comprehension-based model of web navigation and its application to web usability analysis. In S. McDonald, Y. Waern & G. Cockton (Eds.), *People and computers XIV - usability or else!* (Proceedings of HCI 2000) (pp. 357-373). Berlin: Springer.

Kohring, M. & Matthes, J. (2007). Trust in news media: validation of a multidimensional scale. *Communication Research,* 34(2), 231-252.

Koyani S. J., Bailey, R. W. & Nall, J. R. (2004). *Research-based web design and usability guidelines.* Washington, DC: Department of Health & Human Services.

Krahmer, E. & Ummelen, N. (2004). Thinking about thinking aloud: a comparison of two verbal protocols for usability testing. *IEEE Transactions on Professional Communication, 47*(2), 105-117.

Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.

Lavie, T. & Tractinsky, N. (2004). Assessing dimensions of perceived visual aesthetics of web sites. *International Journal of Human-Computer Studies, 60*(3), 269-298.

Law, E., Roto, V., Vermeeren, A., Kort, J. & Hassenzahl, M. (2008). Towards a shared definition of user experience, *CHI 2008 Proceedings*, 2395-2398.

Law, E., Roto, V., Hassenzahl, M., Vermeeren, A. & Kort, J. (2009). Understanding, scoping and defining User eXperience: a survey approach. *Proceedings of the CHI 2009 Conference on Human Factors in Computing Systems*, 719-728.

Law, E. & Schaik, P. van (2010). Modelling user experience - an agenda for research and practice. *Interacting with Computers, 22*(5), 313-322.

Lederer, A. L., Maupin, D. J., Sena, M. P. & Zhuang, Y. (2000). The technology acceptance model and the World Wide Web. *Decision Support Systems, 29*(3), 269-282.

Li, Y.-M. & Yeh, Y.-S. (2010). Increasing trust in mobile commerce through design aesthetics. *Computers in Human Behaviour, 26*(4), 673-684.

Li, X. (2002). Web page design affects news retrieval efficiency. *Newspaper Research Journal, 23*(1), 38-49.

Lindgaard, G., Fernandes, G., Dudek, C. & Brown, J. (2006). Attention web designers: you have 50 milliseconds to make a good first impression! *Behaviour & Information Technology, 25*(2), 115-126.

Lingle, J. H. & Ostrom, T. M. (1979). Retrieval selectivity in memory-based impression judgements. *Journal of Personality and Social Psychology, 37*(2), 180-194.

MacKinnon, D. P., Lockwood, C. M., West, S. G. & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological methods, 7*(1), 83-104.

Magni, M., Taylor, M. S. & Venkatesh, V. (2010). 'To play or not to play': a cross-temporal investigation using hedonic and instrumental perspectives to explain user intentions to explore a technology. *International Journal of Human-Computer Studies, 68*(9), 575-588.

Mahlke, S. & Minge, M. (2008). Consideration of multiple components of emotions in human-technology interaction. In C. Peter & R. Beale (Eds.), *Affect and Emotion in Human-Computer-Interaction* (pp. 51-62). Berlin: Springer.

Marcoulides, G. A. & Saunders, C. (2006). Editor's comments. PLS: a silver bullet? *MIS Quarterly, 30*(2), iii-ix.

Martensen, A. & Grønholdt, L. (2003). Improving library users' perceived quality, satisfaction and loyalty: an integrated measurement and management system. *The Journal of Academic Librarianship, 29*(3), 140-147.

McDonald, S. & Stevenson, R. J. (1998). Effects of text structure and prior knowledge of the learner on navigation in hypertext. *Human Factors, 40*(1), 18-27.

Mehrabian, A. & Russell, J. A. (1974). *An approach to environmental psychology*. Cambridge, MA: MIT Press.

Monk, A. (2004). The product as a fixed-effect fallacy. *Human-Computer Interaction, 19*(4), 371-375.

References

Monsuwé, T.P.Y., Dellaert, B.G.C. & Ruyter, K. de (2004). What drives consumers to shop online? A literature review. *International Journal of Service Industry Management, 15*(1), 102-121.

Moshagen, M., Musch, J. & Göritz, A. (2009). A blessing, not a curse: experimental evidence for beneficial effects of visual aesthetics on performance. *Ergonomics, 52*(10), 1311-1320.

Nah, F. & Davis, S. A. (2002). HCI research issues in electronic commerce. *Journal of Electronic Commerce Research, 3*(3), 98-113.

News Corporation (2009). *Rupert Murdoch before the Federal Trade Commission's Workshop. From town crier to bloggers: how will journalism survive the Internet age?* (2009 December 1). Retrieved May 26, 2011, from http://www.newscorp.com/news/news_435.html

Nguyen, A. (2008). *The penetration of online news*. Saarbrücken , Germany: VDM Verlag.

Nielsen, J. (1993). *Usability engineering*. Boston: Academic Press.

Nielsen, J. (1997). *The need for speed*. Retrieved April 30, 2010, from http://www.useit.com/alertbox/9703a.html

Nielsen, J. (2000). *Designing Web usability: the practice of simplicity*. Thousand Oaks, CA: New Riders Publishing.

Nielsen, J. (2004). *Deceivingly strong information scent costs sales*. Retrieved 26 April, 2010, from http://www.useit.com/alertbox/20040802.html

Nielsen, J. (2009). *IA task failures remain costly*. Retrieved December 8, 2011, from http://www.useit.com/alertbox/ia-failures.html

Nielsen, J. (2012). *Thinking aloud: the #1 usability tool*. Retrieved January 17, 2012, from http://www.useit.com/alertbox/thinking-aloud-tests.html

Norman, D. A. (2005). *Emotional design: why we love (or hate) everyday things*. New York: Basic Books.

Office for National Statistics (2011). *Internet access - households and individuals*. Retrieved 20 December, 2011, from http://www.ons.gov.uk/ons/dcp171778_227158.pdf

Onlinenewspapers.com (2011). *English newspapers (2011, May 26)*. Retrieved December 20, 2011, from http://www.onlinenewspapers.com/englanda-k.htm

Oostendorp, H. van & Juvina, I. (2007). Using a cognitive model to generate web navigation support. International Journal of Human-Computer Studies, 65(10), 887-897.

O'Reilly, T. (2007). What Is Web 2.0: design patterns and business models for the next generation of software. *International Journal of Digital Economics, 65*, 17-37.

Pavlou, P. A. & Gefen, D. (2004). Building effective online marketplaces with institution-based trust. *Information Systems Research, 15*(1)*,* 37-59.

Payne, J. W., Bettman, J. R. & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge, UK: Cambridge University Press.

Pedhazur, E. J. & Pedhazur-Schmelkin, L. (1991). *Measurement, design, and analysis: an integrated approach*. London: Lawrence Erlbaum Associates.

Petrie, H., Hamilton, F., King, N. & Pavan, P. (2006). Remote usability evaluations with disabled people*. CHI 2006 Proceedings*, 1133-1141.

Petrie, H., Papadofragkakis, G., Power, C. & Swallow, D. (2009). Navigational consistency in websites: what does it mean to users? *Lecture Notes in Computer Science (LNCS): Vol. 5726 (Part 1)*, (pp. 423-427).

Pirolli, P. & Card, S. (1999). Information foraging. *Psychological Review, 106*(4), 643-675.

Porat, T. & Tractinsky, N. (in press). It's a pleasure buying here: the effects of web-store design on consumer's emotions and attitudes. *Human-Computer Interaction.*

Porat, T. & Tractinsky, N. (2006). The effects of perceived web-store characteristics on consumers' affective states and attitudes towards the store. *Proceedings of the Fifth Conference on Design&Emotion* 2006, September 27-29, Gothenburg, Sweden.

Porat, T., Liss R., Tractinsky, N. 2007. E-stores design: the influence of e-store design and product type on consumers' emotions and attitudes. 12th International Conference on Human-Computer Interaction (HCI) 2007, July 22-27, Beijing. Published in: *Lecture Notes in Computer Science (LNCS): Vol. 4553* (pp. 712-721). Berlin/Heidelberg: Springer.

Power, C., Petrie, H. & Mitchell, R. (2009). A framework for remote user evaluation of accessibility and usability of websites. *Lecture Notes in Computer Science, (LNCS): Vol. 5614* (pp. 594-601)., Berlin/Heidelberg: Spinger-Verlag.

Redelmeier, D. A. & Kahneman, D. (1996). Patients' memories of painful medical treatments: real-time and retrospective evaluations of two minimally invasive procedures. *Pain, 66*(1), 3-8.

Riffe, D., Lacy, S. & Fico, F. G. (2005). *Analysing media messages: using quantitative content analysis in research* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Rosenfeld, L. & Morville, P. (2006). *Information architecture for the world wide web: designing large-scale web sites* (3rd ed.). Sebastopol, CA: O'Reilly.

Rosenstiel, T. & Mitchell, A. (2011). *The state of the news media, 2011: an annual report on American journalism*. Retrieved May 26, 2011, from http://stateofthemedia.org

Roto, V., Law, E., Vermeeren, A. & Hoonhout, J. (Eds.). (2011). *User experience white paper. Results from the Dagstuhl seminar on demarcating user experience, September 15-18, 2010.* Retrieved November 16, 2008, from http://www.allaboutux.org/files/UX-WhitePaper.pdf

Rowley D. & Rhoades D. (1992). The cognitive jogthrough: a fast-paced user interface evaluation procedure. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 11*, 389-395. Retrieved December 8, 2010, from http://dl.acm.org/citation.cfm?id=142869&bnc=1

Russo, J. E., Johnson, E. & Stephens, D. L. (1989). The validity of verbal protocols. *Memory and Cognition, 17*(6), 759-769.

Ryan, M. J., Buzas, T. & Ramaswamy, V. (1995). Making customer satisfaction measurement a power tool. *Marketing Research, 7*(3), 11-16.

Ryan, R. M. & Deci, E.L. (2000a). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54-67.

Ryan, R. M. & Deci, E. L. (2000b). Self-determination theory and the facilitation of intrinsic motivation, social development and well-being. *American Psychologist, 55*(1), 68-78.

Schaik, P. van, Hassenzahl, M. & Ling, J. (under review). Modelling user-experience from an inference perspective. *Transactions on Computer-Human Interaction.*

Schaik, P. van & Ling, J. (2003). Using on-line surveys to measure three key constructs of the quality of human-computer interaction in web sites:

psychometric properties and implications. *International Journal of Human-Computer Studies, 59*(5), 545-567.

Schaik, P. van & Ling, J. (2007). Design parameters of rating scales for web sites. *ACM Transactions on Computer-Human Interaction, 14*(1), Article 4.

Schaik, P. van & Ling, J. (2008). Modelling user experience with web sites: usability, hedonic value, beauty and goodness. *Interacting with Computers, 20*(3), 419-432.

Schaik, P. van & Ling, J. (2009). The role of context in perceptions of the aesthetics of web pages over time. *International Journal of Human-Computer Studies, 67*(1), 79-89.

Schaik, P. van & Ling, J. (2011). An integrated model of interaction experience for information retrieval in a Web-based encyclopaedia. *Interacting with Computers, 23*(1), 18-32.

Schaik, P. van & Ling, J. (in press [a]). An experimental analysis of experiential and cognitive variables in web navigation. *Human-Computer Interaction.*

Schaik, P. van & Ling, J. (in press [b]). A cognitive-experiential approach to modelling web navigation. *International Journal of Human-Computer Studies*.

Scherer, K. R. (1984). On the nature and function of emotion: A component process approach. In K. R. Scherer & P. Ekman (Eds.), *Approaches to emotion* (pp. 293-317). Hillsdale: Lawrence Erlbaum Associates.

Scherer, K. R. (2001). Appraisal considered as a process of multi-level sequential checking. In: K. R. Scherer, A. Schorr & T. Johnstone (Eds.), *Appraisal processes in emotion: theory, methods, research* (pp. 92-120.). New York: Oxford University Press.

Seneler, C. O., Basoglu, N. & Daim, T. (2009). Exploring the contribution of the design characteristics of information systems' user interface to the adoption process. *International Journal of Business Information Systems,* 4(5), 489-508.

Shackel, B. (1991). Usability - context, framework, definition, design and evaluation. In B. Shackel & S. J. Richardson (Eds.), *Human factors for informatics usability* (pp. 21-31). Cambridge University Press.

Sheldon, K. M., Elliot, A. J., Youngmee, K. & Kasser, T. (2001). What is satisfying about satisfying events? Testing 10 candidate psychological needs. *Journal of Personality and Social Psychology, 80*(2), 325-339.

Smith, P. A. (1996). Towards a practical measure of hypertext usability. *Interacting with Computers, 8*(4), 365-381.

Stone-Romero, E. F. & Anderson, L. E. (1994). Relative power of moderated multiple regression and the comparison of subgroup correlation coefficients for detecting moderating effects. *Journal of Applied Psychology, 79*(3), 354-359.

Sun, H. & Zhang, P. (2008). An exploration of affect factors and their role in user technology acceptance: mediation and causality. *Journal of the American Society for Information Science and Technology, 59*(8), 1-12.

Sutcliffe, A. G. (2002). Assessing the reliability of heuristic evaluation for website attractiveness and usability. *Proceedings of the Hawaii International Conference on System Sciences (HICCS), 35*, 1838-1847.

Sutcliffe, A. G. & De Angeli, A. (2005). Assessing interaction styles in web user interfaces. *Proceedings of Human Computer Interaction (Interact 2005)*, 405-417.

Thüring, M. & Mahlke, S. (2007). Usability, aesthetics and emotions in human–technology interaction. *International Journal of Psychology, 42*(4), 253–264.

Toms, E. (2000). Understanding and facilitating the browsing of electronic text. *International Journal of Human-Computer Studies, 52*,(3), 423-452.

Tractinsky, N., Cokhavi, A., Kirschenbaum, M. & Sharfi, T. (2006). Evaluating the consistency of immediate aesthetic perceptions of web pages. *International Journal of Human-Computer Studies, 64*(11), 1071-1083.

Tractinsky, N., Katz, A. S. & Ikar, D. (2000). What is beautiful is usable. *Interacting with Computers, 13*(2), 127-145.

Utting, K. & Yankelovich, N. (1989). Content and orientation in hypermedia networks. *ACM Transactions on Information Systems, 7*(1), 58-84.

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization*. Perceptual and Motor Skills, 47*(2)*,* 599–604.

Venkatesh, V. (2000). Determinants of perceived ease of use: integrating control, intrinsic motivation, and emotion into the technology acceptance model. *Information Systems Research, 11*(4), 342-365.

Venkatesh, V. & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: four longitudinal field studies. *Management Science, 46*(2), 186-204.

Venkatesh, V., Morris, M. G., Davis, F. D. & Davis, G. B. (2003). User acceptance of information technology: toward a unified view. *MIS Quarterly, 27*(3), 425-478.

Vinzi, V. E., Chin, W. W., Henseler, J. & Wang, H. (Eds.). (2010). *Handbook of partial least squares: concepts, methods and applications in marketing and related fields*. Berlin: Springer.

Wang, Y. D. & Emurian, H. H. (2005). An overview of online trust: concepts, elements, and implications. *Computers in Human Behavior, 21*, 105-125.

Watson, D., Clark, L. A. & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063-1070.

White, R. (1959). Motivation reconsidered: the concept of competence. *Psychological Review, 66*(5), 297-333.

Wirtz, D., Kruger, J., Scollon, C. N. & Diener, E. (2003). What to do on spring break? The role of predicted, on-line, and remembered experience in future choice. *Psychological Science, 14*(5), 520-524.

Wixom, B. H. & Todd, P. A. (2005). A theoretical integration of user satisfaction and technology acceptance. *Information Systems Research, 16*(1), 85-102.

Woods, D. D. (1984). Visual momentum: a concept to improve the cognitive coupling of person and computer. *International Journal of Man-Machine Studies, 21*(3), 229-244.

Yang, Z., Cai, S., Zhou, Z. & Zhou, N. (2005). Development and validation of an instrument to measure user perceived service quality of information presenting Web portals. *Information & Management, 42*(4)*,* 575-589.

Zajonc, R. B. (1980). Feeling and thinking: preferences need no inferences. *American Psychologist, 35*(2), 151-175.

Zhang, P. & Dran, G. M. von (2000). Satisfiers and dissatisfiers: a two-factor model for website design and evaluation. *Journal of the American Society for Information Science, 51*(14), 1253-1268.

Zhou, H. & Fu, X. (2007). Understanding, measuring, and designing user experience: the causal relationship between the aesthetic quality of products and user affect. In J. A. Jacko (Ed.), *Human-computer interaction. Interaction design and usability* (pp. 340-349). Berlin: Springer.

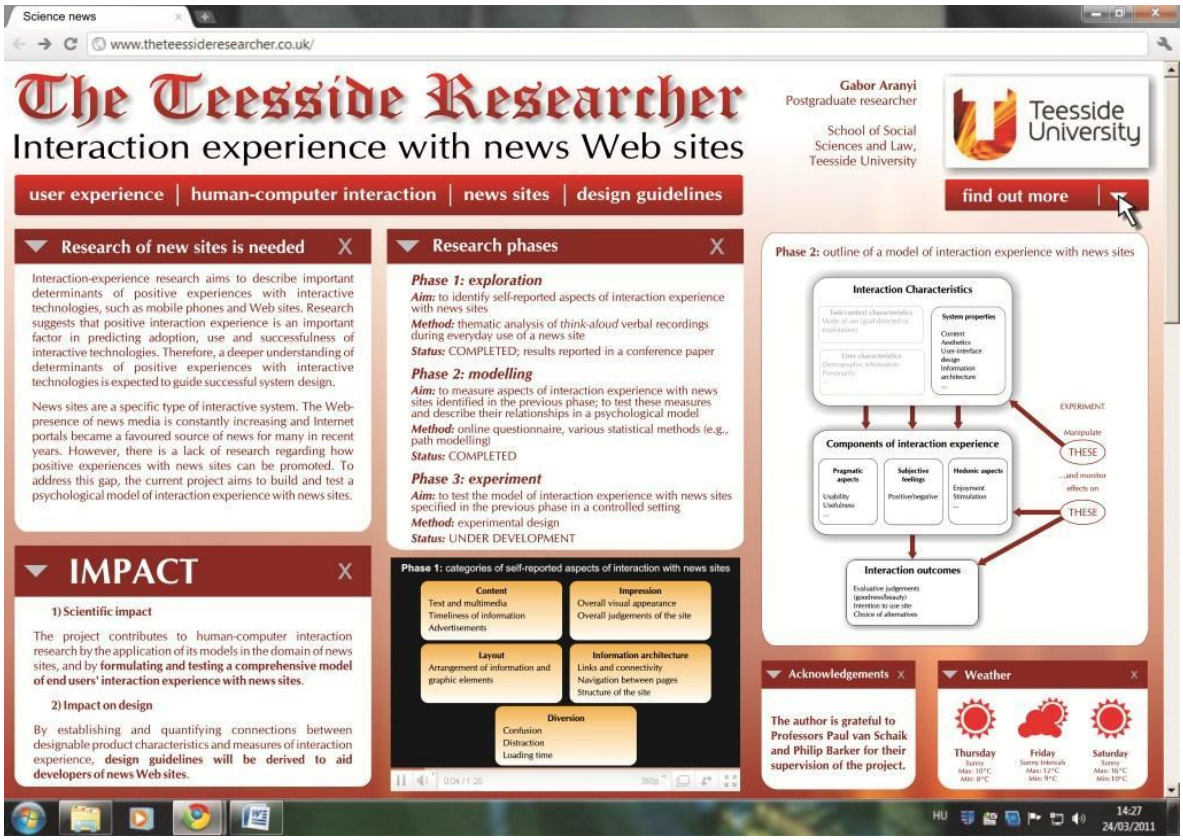
Zijlstra, F.R.H. (1993). *Efficiency in work behavior. A design approach for modern tools*. PhD thesis, Delft University of Technology. Delft, The Netherlands: Delft University Press.

External publication

A condensed version of Chapter 3 with the title *Using think-aloud and psychometrics to explore users' experiences with a news website* was submitted to the British HCI 2011 Conference. The manuscript was published by the journal *Interacting with Computers* (Aranyi et al., 2012).

Poster

A poster entitled *Interaction experience with news Web sites* was presented at the Annual University Postgraduate Conference at Teesside University on 24 May 2011 and won the 1st prize of the conference's poster competition. A picture of the poster is presented below.



Appendices

Appendix 2.1. A summary of studies related to the interaction-experience models presented in Chapter 2

Source	Artefact	Task	Measures of experience	Person measures	Artefact measures	Task/context measures	Objective measures	Interaction outcomes
Components of user experience (CUE) model								
Thüring and Mahlke (2007), Study 1	Mobile-phone computer simulations	Navigation	Affective reactions (valence and arousal) {D} Emotional responses (cognitive appraisals)	N/A	Usability [high/low]	N/A	Heart rate {D} Electrodermal activity {D} Electromyographic activity {D} Effectiveness	N/A
Thüring and Mahlke (2007), Study 2	Portable-audio-player computer simulations	Navigation	N/A	N/A	Usability [high/low] Aesthetics [high/low]	N/A	Task completion rate Time on task	Overall rating
Thüring and Mahlke (2007), Study 3	Portable-audio-player computer simulations	Navigation	Affective reactions (valence and arousal) {D}	N/A	Usability [high/low] Aesthetics [high/low] {B}	N/A	Heart rate {D} Electrodermal activity {D} Electromyographic activity {D}	Overall rating Preference
Model of user-interface quality assessment								
Hartmann et al. (2008), Study 1 [Experiments 1 and 2, reported in Sutcliffe & De Angeli, 2005, and De Angeli et al., 2006, respectively]	Two educational Web sites	Information retrieval	Engagement	Target population (children/adults)	Interaction style [menu/metaphor] Aesthetics (classical/expressive) Memorability Usability {D,A} Information and service quality	Scenario of use (serious/leisure)	Usability (number of errors) Memory recall	Overall preference
Hartmann et al. (2008), Study 2 (Experiment 3)	Mobile phone	Reading	Look and feel Engagement	N/A	Content-fit [generic/customised] Aesthetics [high/low] (classical/expressive) Content and service quality	Mobility [idle/mobile]	Reading time	Overall preference

Appendix 2.1

Source	Artefact	Task	Measures of experience	Person measures	Artefact measures	Task/context measures	Objective measures	Interaction outcomes
Hartmann et al. (2007)	Three departmental Web sites of a university	Interview (during browsing)	N/A	Users' background (cultural and training)	Usability Aesthetics Content quality Reputation/identity	Scenario of use (critical/less critical)	N/A	Overall preference
Environmental-psychology model								
Porat and Tractinsky (in press)	4 online-retail sites (2 domains)	Browsing and simulated purchase	Affect (pleasure, arousal and dominance)	N/A	Usability Aesthetics (classical/expressive)	N/A	N/A	Attitudes towards stores
User-experience (UX) model								
Hassenzahl et al. (2002)	2 Web sites	Exploration or information search	N/A	N/A	Pragmatic quality Hedonic quality	Mode of use (goal/activity)	N/A	Product appeal
Hassenzahl (2004), Study 1	4 MP3-player skins (ugly/beautiful)	Observation (no use)	N/A	N/A	Hedonic quality (stimulation and identification) Pragmatic quality	N/A	N/A	Beauty Goodness
Hassenzahl (2004), Study 2	4 MP3-player skins (ugly/beautiful)	Use tasks	Mental effort {D}	N/A	Hedonic quality (stimulation and identification) {B,A} Pragmatic quality {B,A}	N/A	N/A	Beauty {B,A} Goodness {B,A}
Hassenzahl and Ullrich (2007)	An interactive storytelling tool	Free exploration or information search	Mental effort {D} Affect (valence) {D} Spontaneity	N/A	N/A	Mode of use [action/goal]	Knowledge acquisition	Product appeal
Hassenzahl (2008)	Self-reported selection of technologies	Recall of satisfying use episode (no use)	Need fulfilment (autonomy, competence and relatedness) Affect (positive/negative)	N/A	N/A	N/A	N/A	N/A
Hassenzahl et al. (2010)	Self-reported selection of technologies	Recall of satisfying use episode (no use)	Need fulfilment (7 needs) Positive affect	N/A	Pragmatic quality Hedonic quality	N/A	N/A	N/A

Appendix 2.1

Source	Artefact	Task	Measures of experience	Person measures	Artefact measures	Task/context measures	Objective measures	Interaction outcomes
Hassenzahl and Monk (2010)	Web sites	Browsing and evaluation	N/A	N/A	Pragmatic quality Hedonic quality	N/A	N/A	Beauty Goodness
Van Schaik et al. (under review), Study 1	An online encyclopaedia	Observation and exploration	N/A	N/A	Pragmatic quality {B,A} Hedonic quality {B,A}	N/A (action mode only)	N/A	Beauty {B,A} Goodness {B,A}
Van Schaik et al. (under review), Study 2	A council Web site	Observation and information retrieval	N/A	N/A	Pragmatic quality {B,A} Hedonic quality {B,A}	N/A (goal mode only)	N/A	Beauty {B,A} Goodness {B,A}
Van Schaik et al. (under review), Study 3	A bespoke university intranet site	Observation and information retrieval	Perceived challenge {D}	N/A	Artefact complexity [low/high] Pragmatic quality {B,A} Hedonic quality {B,A}	Task complexity [low/high] (goal mode only)	N/A	Beauty {B,A} Goodness {B,A}
Van Schaik and Ling (2008)	An intranet site	Observation and set tasks	Mental effort {D}	N/A	Pragmatic quality {B,A} Hedonic quality (stimulation and identification) {B,A} Screen design [good/bad] Information organisation [good/bad]	N/A	Task performance	Beauty {B,A} Goodness {B,A}
Van Schaik and Ling (2009), Study 1	Screenshots of 50 home pages of local-government sites	Observation (500ms and self-paced exposure) (no use)	N/A	N/A	Attractiveness [high/low] Aesthetics (classical/expressive)	Mode of use [action, goal and control]	N/A	N/A
Van Schaik and Ling (2009), Study 2	A bespoke university Web site	Observation (500ms and self-paced exposure) and use	Mental effort {D}	N/A	Design aesthetics [classical and expressive] {B,A} Attractiveness {B}	Mode of use [action, goal and control]	Task performance	N/A

Appendix 2.1

Source	Artefact	Task	Measures of experience	Person measures	Artefact measures	Task/context measures	Objective measures	Interaction outcomes
Van Schaik and Ling (2011a)	An online encyclopaedia	Viewing and use (information retrieval or exploration)	Perceived enjoyment	N/A	Usability [high/low] Perceived usefulness Perceived ease of use Pragmatic quality {B,A} Hedonic quality (identification and stimulation) {B,A}	Mode of use [goal/action]	Task performance	Intention to use Beauty {B,A} Goodness {B,A}

Note. Brackets indicate dimensions of the applied measures. Square brackets under the headings of types of measures indicate experimental manipulation of the particular characteristic. Psychological measures applied to measure the manipulated characteristics are not listed separately. For those studies that included interaction with the artefacts, letters in curly brackets indicate times of measurement (B: before interaction, D: during interaction and A: after interaction). Measurement after interaction is treated as the default; therefore, it is only indicated when another time of measurement is also present for a particular measure.

Appendix 3.1. Online questionnaire (Study 1)

FIRST WEB PAGE

[INTRODUCTION]

Welcome

The University of Teesside is conducting a study to explore how people use online news sites. We would like to ask you to participate in this short survey.

We need people who use www.gazettelive.co.uk regularly, who are over 18, and who speak fluent English.

Some of those who complete this survey will be invited to take part in a practical exercise at the University of Teesside. If you wish to take part in the exercise, please leave your e-mail address at the end of the questions.

All data obtained in the study will be treated confidentially. The researcher will not give out your e-mail address to a third person.

Participants will remain anonymous. Therefore please choose a nickname/pseudonym for yourself which you will be able to recall later.

SECOND WEB PAGE

[DEMOGRAPHICS]

1. Your nickname/pseudonym: [Text box]

2. How old are you (years)? [Text box]

3. Are you? [Radio buttons]

Male
Female

4. What is the highest level of education that you have achieved? (Please select one answer.) [Radio buttons]

O level/GCSE or equivalent
NVQ or equivalent
A level or equivalent

TROCN/Btec Dip
Degree
Higher degree
No qualification
Other [Text box]

5. What is your situation in relation to work? (Please select one answer.) [Radio buttons]

Retired
Employed
Self-employed
Student
Full-time homemaker
Unemployed/looking for work
Unfit to work
No response/rather not say
Other [Text box]

[INTERNET USE BEHAVIOUR]

6. How long have you been using the Internet (years)? [Text box]

7. Do you have Internet connection at home? [Radio buttons]

Yes
No

8. Where do you access the Internet? (Select all that apply.) [Checkboxes]

Home
Work
Library
Internet café
Other [Text box]

9. What kind of devices do you use to access the Internet? (Select all that apply.) [Checkboxes]

Desktop PC
Laptop PC
PDA
Mobile Phone

Other [Text box]

10. How often do you log on to the Internet? (Please select one answer.) [Radio buttons]

Weekly
2-3 times a week
4-6 times a week
Daily
2-3 times a day
More than 3 times a day

11. Once on line, how much time do you spend on the Internet on average? (Please select one answer.) [Radio buttons]

1-5 minutes
6-10 minutes
About 15 minutes
About 30 minutes
About 45 minutes
About 1 hour
Several hours

12. What do you use the Internet for? (Select all that apply.) [Checkboxes]

E-mail
Reading news
Buying products or services (e.g., general shopping, train tickets, books, insurance and travel)
Social networking (e.g., Facebook and YouTube)
Visiting chat rooms, forums
Managing bank accounts
Searching for work-related or study-related information
Education/training
Surfing (exploring the Web)
Downloading or file sharing (e.g., BitTorrent)
Working

13. Your most frequently visited Web sites. (Please type the name of one or more Web sites that you visit.)

Which Web site do you visit most frequently? [Text box]
What is your second-most frequently visited Web site? [Text box]

What is your third-most frequently visited Web site? [Text box]
What is your fourth-most frequently visited Web site? [Text box]
What is your fifth-most frequently visited Web site? [Text box]

[NEWS SITE USE BEHAVIOUR]

14. How often do you visit online news sites? (Please select one answer.) [Radio buttons]

Weekly
2-3 times a week
4-6 times a week
Daily
2-3 times a day
More than 3 times a day

15. How much time do you spend on online news sites? (Please select one answer.) [Radio buttons]

1 minute
2-4 minutes
About 5 minutes
About 10 minutes
About 15 minutes
About 30 minutes
More

16. Which online news site(s) do you visit regularly? (Please type the name of one or more news sites that you visit.)

News Site 1 [Text box]
News Site 2 [Text box]
News Site 3 [Text box]
News Site 4 [Text box]
News Site 5 [Text box]

17. Do you comment the news and blog entries on news sites? (Please select one answer.) [Radio buttons]

Never
Incidentally
Sometimes
Often

Very frequently

18. Do you prefer online news sites over printed newspapers? [Radio buttons]

Yes
No

19. Do you purchase printed newspapers regularly or are you subscribed to any daily papers? [Radio buttons]

Yes
No

[CONTACT INFORMATION]

20. Are you interested in taking part in a practical exercise (30-45 minutes long) at the University of Teesside? [Radio buttons]

Yes
No

21. If you are interested in taking part, please provide your e-mail address below to allow the researcher to contact you (optional). [Text box]

THIRD WEB PAGE

Done

Thank you for completing this survey.
To navigate back to Gazette, click the link below (new window opens).
http://gazettelive.co.uk

Appendix 3.2. Participant information sheet, consent form and post-use interaction-experience questionnaires (Study 1)

Participant Information Sheet

Project title: developing a psychological model of end-user's interaction with news sites

Dear Madam or Sir,

Thank you for participating. The purpose of this study is to explore how people use online news sites. You will be asked to use a site and express your thoughts and feelings about it and to answer a series of questions regarding your Internet usage habits and experience with news sites. With your help, we aim to have a better understanding of what makes a good news site and how people express their opinion about Web sites. It is expected that the results of the research will help in the design of better Web sites. The procedure will take approximately 1 hour of your time.

Your personal details and the data obtained during the study will be stored separately and treated confidentially. Data will be stored in a locked filing cabinet in a university building and will only be accessed by the researchers. Data will be kept until the completion of the research.

Your comments and actions during the interaction with the site will be recorded and the recording will be stored securely on a compact disk in a locked filing cabinet until the completion of the research.

Your anonymity will be maintained during the whole research. An identification number rather than your name will be used to identify your data and no personal information or individual cases will be published.

If you feel uncomfortable with the procedure, you have the right to withdraw at any time during the research without any personal consequences, up to the start of data analysis.

If you have any queries about the purpose and procedure of the study, please feel free to ask the researcher. Once again, thank you for taking part in this study.

[NAME AND CONTACT DETAILS OF THE RESEARCHER]

Appendix 3.2

Informed Consent
Consent form

Teesside University
School of Social Sciences and Law

Project title: developing a psychological model of end-user's interaction with news sites

I confirm, that (please tick)

I have been informed of the purpose of the study as described in the Participant Information Sheet.
I have been given the opportunity to ask questions about the study.
I have the right to withdraw at any time - without personal consequences - up to the start of data analysis.
My personal details and the data obtained during the study will be stored separately and treated confidentially. Data will be stored in a locked filing cabinet in a university building and will only be accessed by the researchers. Data will be kept until the completion of the research.
I understand that my behaviour will be recorded and the recording will be stored securely on a compact disk in a locked filing cabinet until the completion of the research.
I agree to participate in the study as outlined to me.

.....	.....	.....
Name of participant	Date	Signature
.....	.....	.....
Name of witness	Date	Signature

Researcher: Gabor Aranyi

School of Social Sciences and Law
Teesside University
TS1 3BA Middlesbrough, UK
E-mail: h8120322@tees.ac.uk

Appendix 3.2

Post-use questionnaires

Perceived enjoyment (Sun and Zhang, 2008)

Please recall your use of the web site. You will find ten words below regarding the look and design of the site. Please indicate to which extent you find the words characteristic of the site by drawing a circle around the number of your choice. Please keep the meanings of the following possible answers in mind as you rate the appropriateness of the words describing the site:

1= strongly disagree
4 = neutral
7= strongly agree

	Strongly disagree			Neutral			Strongly agree
1. I find using the Gazette to be enjoyable.	1	2	3	4	5	6	7
2. The actual process of using the Gazette is pleasant.	1	2	3	4	5	6	7
3. I have fun using the Gazette.	1	2	3	4	5	6	7

Intensity of flow (Davis & Wiedenbeck, 2001)

Please recall your use of the web site. You will find nine statements below. Please indicate to which extent you agree with each statement by drawing a circle around the number of your choice. Please keep the following scale in mind as you rate each of the statements below:

1= strongly disagree
4 = neutral
7= strongly agree

Appendix 3.2

	Strongly disagree			Neutral			Strongly agree
1. I thought about other things.	1	2	3	4	5	6	7
2. I had to make an effort to keep my mind on the activity.	1	2	3	4	5	6	7
3. I was aware of distractions.	1	2	3	4	5	6	7
4. I was aware of other problems.	1	2	3	4	5	6	7
5. Time seemed to pass more quickly.	1	2	3	4	5	6	7
6. I knew the right things to do.	1	2	3	4	5	6	7
7. I felt like I received a lot of direct feedback.	1	2	3	4	5	6	7
8. I felt in control of myself	1	2	3	4	5	6	7
9. I felt in harmony with the environment.	1	2	3	4	5	6	7

Perceived aesthetics (Lavie & Tractinsky, 2004)

Please recall your use of the web site. You will find ten words below regarding the look and design of the site. Please indicate to which extent you find the words characteristic of the site by drawing a circle around the number of your choice. Please keep the meaning of the following possible answers in mind as you rate the appropriateness of the words describing the site:

1= strongly disagree
4 = neutral
7= strongly agree

	Strongly disagree			Neutral			Strongly agree
Original	1	2	3	4	5	6	7
Clean	1	2	3	4	5	6	7
Sophisticated	1	2	3	4	5	6	7
Clear	1	2	3	4	5	6	7
Fascinating	1	2	3	4	5	6	7
Creative	1	2	3	4	5	6	7
Uses special effects	1	2	3	4	5	6	7
Pleasant	1	2	3	4	5	6	7
Symmetrical	1	2	3	4	5	6	7
Aesthetic	1	2	3	4	5	6	7

Appendix 3.2

AttrakDiff2 abridged (Hassenzahl & Monk, 2010)

Please recall your use of the web site. You will find word pairs below to assist you in the evaluation of the site. Each pair represents extreme contrasts. The possibilities between the extremes enable you to describe the intensity of the quality you choose. Please indicate your choice for each word pair by ticking one of the circles between the words.

Confusing	O	O	O	O	O	O	O	Structured
Unpredictable	O	O	O	O	O	O	O	Predictable
Impractical	O	O	O	O	O	O	O	Practical
Complicated	O	O	O	O	O	O	O	Simple
Dull	O	O	O	O	O	O	O	Captivating
Tacky	O	O	O	O	O	O	O	Stylish
Cheap	O	O	O	O	O	O	O	Premium
Unimaginative	O	O	O	O	O	O	O	Creative
Ugly	O	O	O	O	O	O	O	Beautiful
Bad	O	O	O	O	O	O	O	Good

Appendix 3.3. Description of categories of units of thought, instructions and coding table sent to the coders

Description of categories and sub-categories sent to the coders

Units of thought are defined as expressions of participants' opinion or judgement about various aspects of the site (e.g., content and presentation), or expressions of their experience.

Note. Categories that have sub-categories are defined as the sum of sub-category descriptions. Possible examples are hypothetical, not examples from the actual data.

Impression (category)

This category includes first impressions, overall visual impressions and overall evaluative judgements of the site. The time codes will help to decide whether a unit is a first impression or not. Overall visual impressions mean statements about the visual appearance of the site in general. Overall evaluative judgements mean an evaluative statement about the site as a whole (possible examples: it looks good; it is easy to use; it is a bad/ugly site).

Layout (category)

This category includes comments on the layout of text and graphic elements on the site and article presentation style. For example, how the text is broken up into paragraphs and pages, where certain graphic elements like advertisements and navigation controls are located and how individual pages are presented. (Possible examples: the heading should be bigger, I can hardly read it; it is good to have these paragraphs separated, it reads quite well this way.)

Content (category)

Information content (sub-category)

This sub-category includes comments on site and article content (or lack of) (text and multimedia). Every unit regarding content comes here. They can be comments on the content of the site in general, or on the content of individual articles as well. Units regarding text, picture and video content all fall into this category. (Possible examples: I like that picture, it's nice; I think the author has missed the point here.)

Outdated content (sub-category)

Use this sub-category when a participant identifies a content element (text, picture and video) as outdated (possible example: it was ages ago, why is it here?). If you put a unit here, do not put it in the information content sub-category too.

Advertisement (sub-category)

Advertisements count as content as well, so please categorise units regarding advertisements here. It is expected that this sub-category will overlap with the *distraction* sub-category, so if a unit regarding advertisements are reported to be distracting, you may assign them into both sub-categories.

Information architecture (category)

Links (sub-category)

This includes units regarding hyperlinks and connectivity of the site with other Web sites. (Possible example: there's a lot of useful links here.)

Navigation (sub-category)

This includes units regarding navigation between pages of the site and comments on the search functionality. (Possible example: I find the top navigation tabs well-phrased and useful.)

Structure (sub-category)

This includes units regarding information organisation (or lack thereof, e.g., information overload) on the site (not within individual articles). Assign any unit to this sub-category that is about how information is organised or how it should be organised or categorised. Also put units regarding the lack of proper information organisation here (possible example: everything is at one place, it's confusing – this would also go to the *confusion* sub-category; there is too much information on this page; Why is it broken-up into two pages?).

Diversion (category)

Distraction (sub-category)

Use this sub-category whenever a participant suggests being distracted by something. (Possible examples: these adverts are awfully off-putting; that flashing thing keeps drawing my attention away from the story.)

Confusion (sub-category)

Use this sub-category whenever a participant suggests being confused by something. (Possible examples: now I don't know where exactly I am; there's a lot of information crammed together in a small place, it's quite confusing.)

Loading time (sub-category)

Use this sub-category when a participant makes a comment about loading time. (Possible examples: is it going to take forever? It loads quite quickly, that's good.)

Unassigned (category)

You will not be able to assign every unit into the categories defined above. For example, units were also collected regarding general use habits, but it was decided later not to include them as a separate category. Use this category only if you really cannot put a unit anywhere else. The majority of the units should be able to be assigned to at least one of the categories.

Instructions

You may assign units into multiple categories and sub-categories. For example, a unit regarding distracting adverts on the site may fall into *Content → Advert* and *Diversion → Distraction* at the same time.

Write the numbers of the units into the *unit numbers* column in the table. Please separate unit numbers with a colon, followed by a space (e.g., 2, 138, 141).

For categories that have sub-categories, please write the number of the unit behind the category and the sub-category as well.

For units that could not be assigned to any of the categories, please use the *unassigned* category.

The units start with a number in square brackets and then a time follows, indicating when the unit was made. This may be useful for coding a unit as a first impression. After the units, in square brackets, the place is indicated on the site where the comment was made.

Appendix 3.3

Coding table sent to the coders

ID of coder:

Category	Sub-category	Short description	Unit numbers
Impression		First impressions, overall visual impressions and overall evaluative judgements of the site.	
Layout		Comments on the layout of text and graphic elements on the site and article presentation style.	
Content			
	Information content	Comments on site and article content (or lack of) (text and multimedia).	
	Outdated content	Comments on content being out of date.	
	Advertisement	Comments on advertisements.	
Information architecture			
	Links	Comments regarding hyperlinks and connectivity of the site with other Web sites.	
	Navigation	Comments on navigation between pages of the site and search.	
	Structure	Comments on information organisation (or lack of, e.g., information overload) on the site (not within individual articles).	
Diversion			
	Distraction	Comments suggesting being distracted.	
	Confusion	Comments suggesting confusion.	
	Loading time	Comments on loading time	
Unassigned		Comments that could not be assigned in the categories above.	

Appendix 3.4. Units of thought of the think-aloud study

Note. Units of thought are arranged by categories. Units of thought on which all three coders agreed on to belong to a particular category are indicated with bold numbers. Plain numbers represent units of thought agreed upon by Coders 1 and 3. Each unit starts with a time code that indicates when the unit was left by a participant. The location on the site a participant visited while expressing a unit of thought is indicated in square brackets. Following the location code in square brackets are the group identifier (G: Gazette reader; N: non-Gazette reader), the participant number and the valence rating (+ positive; - negative: n: neutral).

Impression	
1	00:00 “It’s a bright site.” [Home] G2n
24	00:17 “This news site doesn’t look... doesn’t have a very good overview for me, so I’m not sure that I will use it again... [Home] N5-
27	00:43 “yeah, it’s very colourful, therefore... which makes it very confusing to me. Yeah, I’ve mentioned that I see this site the first time in my life. [Home] N5-
68	00:06 “...it looks quite reasonably interesting.” [Home] N8+
70	00:22 “Lots of information here... yeah, the information here is lots. Which I like about Web sites.” [Home] N8+
79	00:01 “I’ve been on this Web site before and I don’t like it, because... I like to use it, because it’s a local Web site, but I don’t like the... I don’t like the content, apart from the content that is interesting for me because it’s local.” [Home] N9-
97	00:06 “My first reaction’s there’s an awful lot of information and I’m really not sure where I’d go first.” [Home] N12-
107	03:28 “This is very similar to BT internet, this particular part where you can click through the news feeds. But again, it’s very, very small and there’s so much information, I... it’s not a very... I don’t think it’s a very good Web site. There’s far too much information. I prefer a Web site that... there’s just one or two things and then you link, either from the left side or the right hand side.” [Home] N12-
112	00:00 “Lots of things are going on.” [Home] N13n
118	02:01 “Still have a look to see if there’s anything interesting. I’m actually thinking to myself this isn’t a very interesting Web site.” [/TS17/] N14-
152	00:07 “First impression is kind a the advertisements are probably distracting away from the news headlines, ‘cause they’ve got the same thing in two places, noticing that motion [...] BBC and things like that try not to capture attention on things, but these things are relatively small in relation to the rest of the page, thinking about it.” [Home] G20-
167	00:06 “It seems quite busy, there’s lots, there’s lots of advertisements, lots of things going on.” [Home] N22n

177 00:00 “This is a nice site. It’s not as bad as it used to be four years ago, when I think it was the last time I visited the Gazette site. I usually buy the Gazette, every now and then.” [Home] N24+

179 00:05 “At least it’s very colourful, not like the other news once I’ve been on.” [Home] N25+

182 02:45 “I don’t like this Web site” [/news/Teesside-news/an article] N25-

190 08:36 “I don’t like this Web site at all.” [Home] N25-

Content

2 06:18 “It’s all antisocial behaviour or drugs.” [/news/Teesside-news/] G2n

3 08:50 “...I don’t even read full story, because it’s all the same.” [/news/Teesside-news/] G2-

4 10:01 “Yes, it’s all swine flu, Corus and drugs.” [Home] G2n

5 00:15 “It’s fairly useful, because it has the today’s news from the Evening Gazette.” [Home] G3+

10 03:38 “...and, what you also notice about the site is that... say, quarter of the stories on the... are about criminal offences, most of them involving violence within the local area...” [/news/Teesside-news] G3n

11 04:27 “What is unusual today is that given the time... twelve minutes past two in the afternoon, is that they’ve not managed to fill yet the whole of the front page with today’s stories. There is still five from Saturday. Of course they don’t publish on Sunday, so there’s none from then. Which is odd, given the... it was probably updated on Saturday about this time and you’d have thought in 48 hours there might have been plenty of news, but it appears there isn’t, or if there is, they haven’t got around to telling this about it yet. And yet, they got around to publishing a paper, because the paper version has been out for about three hours, which presumably is full... can’t think they published it with blank pages. So, things just haven’t got on.” [/news/Teesside-news] G3-

12 06:22 “But what you notice also on this is quite often the so-called latest news is from the day before.” [Home] G3-

16 02:58 “There is a nice big picture in the middle of the article that I’ve noticed straight away. Oh, it’s like a gallery so I can scroll through, see the different pictures. There’re some good photos. [...] Excellent photos of people on bikes...” [/news/Teesside-news/an article] G4+

21 11:35 “Volunteers week 2007, so obviously some of the information here is not quite right, so... maybe they’ve copied and pasted it from last year.” [/news/tees-and-tees-valley/redcar-and-cleveland-focus/an article] G4-

29 01:44 “...okay, that’s regional news, but I don’t see any international news and that looks to me not very... not like a news where... that I can really trust. [Home] N5-

30 01:54 “And there’s too much advertising on it, which also makes it... gives me the feeling that I can’t really trust a Web site, because I think most... very... well-recognised news sites don’t have much advertising on their Web site. [Home] N5-

32 02:42 “Okay, weather forecast on the first site... first page, which is good I think.” [Home] N5+

35 06:18 “So they offer everything and nothing.” [Home] N5n

39 00:44 “I don’t know many places around Teesside, but it’s nice to get to know the area I suppose. [The article contains a description of the area the news happened.]” [/news/Teesside-breaking-news] N6+

40 01:15 “That advert for Lidl keeps distracting me on the right hand side; I keep wanting to look at that.” [/news/Teesside-and-tees-valley/local-regeneration/] N6-

48 04:51 “They’ve even got advertisements for advertisements on the bottom.” [/useful-information] N6n

49 05:10 “...it seems like they’re really to the point though. I must admit, it’s quite good.” [/useful-information/finance/an article] N6+

50 05:17 “...and I don’t know if it’s... depends what you click on, but it seems like the animations [advert] are changing depending on what you click. It might just be a coincidence that there’s a finance one come up.” [/useful-information/finance/an article] N6n

51 05:39 “The links all seem to be adverts as well.” [/useful-information/finance/an article] N6n

56 09:32 “Oh, that’s good.” [The meaning of ‘tags’ explained] N6+

60 01:13 “...quite interesting, quite a lot there.” [/boro-fc/boro-fc-news/an article; comments] G7+

62 02:55 “There is a lot there to read. I think I’d read it if I was on a coffee break or a tea break and sit down and just read it to myself and think about it and enjoy the cynicism about... be hoping that the football team can really get back on track. And there’s more comments here, I’d read those as well and possibly leave a comment myself, but not right now, ‘cause... I don’t feel like it.” [boro blogs] G7n

63 03:29 “Something here about business and economy news, quite interesting, but I get a lot of this at the TV, I wake up early and I watch the BBC breakfast news and then I watch news in the evening so I get a lot of that anyway...” [Home] G7+

64 04:09 “Some interesting things here, amusing to look at now, look at their clothes and their style and the way the world was. I find this sort of thing interesting.” [/news/north-east-history/galleries/a gallery] G7+

65 04:40 “This is quite interesting... local interest, history, I find it quite fascinating.” [/news/north-east-history/galleries/a gallery] G7+

72 01:40 “But anyway, that’s all the information there that you could ever want about your local football team in the area [/boro-fc] N8+

73 01:50 “Hmm, what have we got here? Hot deals. This looks interesting.” [/boro-fc/; advert bar] N8+

74 05:10 “So you could ask is this information still... still relevant.” [/news/sirf/] N8-

77 10:10 “That looks like the same photograph we’ve just seen. Anyway...”

[/north-east-sport/motorsport-news/redcar-bears/an article] N8-

78 10:52“At the minute we don’t see the half, the result, which is slightly disappointing, but never mind.” [/north-east-sport/motorsport-news/redcar-bears/an article] N8-

81 09:05 “There are always men on all the photos and no women.” [/north-east-history/galleries/a gallery] N9n

82 00:13 “It’s not really proper news. It seems like quite personal news. To say, someone has had a wedding.” [Home] N10-

83 01:50 “It’s not many pictures of the games.” [/boro-fc/an article] N10-

84 03:02 “Doesn’t seem to be much actual news on it.” [Home] N10-

85 03:46 “It’s not very up to date I think.” [/tags/corus] N10-

87 09:13 “It’s good that there was a video on it, ‘cause I didn’t know it was going on...” [/videos-pictures/news-photos-videos/a video] N10+

88 9:50 “Again, more adverts. [...] It’s kind of the first you see when you look at it. There’s adverts about giving in your gold for like, money... pawnbrokers. It seems like they just preying on people ‘cause they know they’ve got less money now because there is a recession.” [Home] N10-

101 01:14 “Just linking on to sports, there’s an awful lot of information here. There’s a lot of... flash and links at the top... about the army. There’s a lot of other links down the side that’s quite off-putting, actually.” [/north-east-sport/] N12-

102 02:00 “The latest update is 22 September, which is quite old. There’s nothing new on there, that’s not very good for me.” [/north-east-sport/latest-sport-news/an article] N12-

114 04:33 “The first thing I notice is the army advert again. The adverts are very prominent on the top, it’s a good place to advertise, looks like. And a little film there... [An advert video playing on the right]” [/north-east-sport/] N13n

117 08:30 “...petrol prices, that’s good.” [/motors/]N13+

127 06:11 “...but no picture with that [article]. [...] I think it’d be better if it was more like a proper newspaper that... it was, there was pictures and text. There were pictures and text but then, like you didn’t have the adverts. And those moving adverts... really distracting. I don’t like them at all. But this has no pictures at all [the article]. But at least it’s all on one page this time [...] ” [/north-east-sport/north-east-cricket/an article] N15-

133 01:27 “When I look at it... So this is breaking news. There’s actually very little there. So... it doesn’t really tell me much at all. It tells me that there is a full report in tonight’s Evening Gazette. I’m presuming it again it doesn’t tell me if it’s in the paper itself, which I don’t buy, or whether it will be on here tonight. So I don’t think it’s really helpful, plus it doesn’t kind of interpret the crime figures for me.” [/news-Teesside-breaking-news/an article] N16-

134 02:10 “I suppose one of the things I think about web... these sorts of Web sites are the kind of number of stories that don’t really have much to say. Which sounds awful, I realise that. [Looks at a latest story] Like someone... a team

wearing pink as part of a cancer appeal. I second the cancer appeal is very important wearing pink but it's not really for me headline news." [Home] N16-

135 03:25 "There's lots of flashing things. I don't like advertising, but I realise that local... some... well, a lot of Web sites rely on that. There's a lot of red on it, which is quite... reminds me of something actually. I don't know what it is. It seems a bit too busy for me." [TS12/an article] N16-

136 03:59 "Again, it kind of doesn't really tell me anything, but I think that might be because I don't live here. I'm not part of the community in this area..." [TS12/an article] N16-

137 05:23 "I like... Yes, something like this is kind of interesting, 'cause I like stuff about local identity. And this is about eating parmos, which is a phenomenon I never knew about until I started working here." [A blog entry] N16+

139 08:00 "The advertising is really annoying. It's actually quite big on the screen and so... I'm not used to that... I'm used to two smaller sidebars really, but that's quite distracting. Oh, I'll just ignore it." [/videos-pictures/your-uploaded-pictures/] N16-

140 08:25 "The latest image gallery is March 2009, which doesn't appear very latest to me." [/videos-pictures/your-uploaded-pictures/] N16-

143 00:37 "There isn't much information on cricket. There's no update on the cricket match." [/north-east-sport/latest-sport-news/an article] N17-

144 07:23 "What is this telling, it was ages ago. Yeah, where is the actual recent results from property?" [/news/Teesside-news/an article] N18-

147 02:24 "Quite a few adverts on the right. Don't really like how they seem to flash up all over the page. They tend to make the page quite slow." [Home] G19-

151 07:55 "Again, there's quite a few flashing ads and big, colourful things flashing all over the place, which tends to draw your eye away from what you... what you really looking at." [/news/.../Stockton-focus/an article] G19-

152 00:07 "First impression is kind a the advertisements are probably distracting away from the news headlines, 'cause they've got the same thing in two places, noticing that motion [...] BBC and things like that try not to capture attention on things, but these things are relatively small in relation to the rest of the page, thinking about it." [Home] G20n

153 00:47 "Scrolling down to see some of the main headlines, but it's kind a very local orientated and things rather than... generic things for the area. But it does obviously offer the option to be bit more specific... [Home] G20n

156 02:17 "Now I've just noticed, see the latest news scrolling past [under latest news pictures], but kind of like a typewriter, which is very slow and very... doesn't really give you much information about anything that's going on there." [Home] G20-

161 04:35 "Kind of good to see really that it doesn't just... football just doesn't dominate it." [/sport/] G20+

162 04:58 "Again, probably a bit more generic and probably nicely focused

towards the area, rather than the national standard of just football, cricket, rugby and things like that. [/sport/] G20+

165 08:11 “The most recommended [...] it’s all again very local orientated, which is good in some ways, but if you’re not interested in that area it’s not really kind of your thing.” [Home] G20+

167 00:06 “It seems quite busy, there’s lots, there’s lots of advertisements, lots of things going on.” [Home] N22n

168 01:13 “The news are laid out well, but there’s a lots of sponsored links on the right-hand side of the page advertising jobs and like... almost like a local business finder, which I’m sure would be very good if I was looking for work, but I think it kind a negates what I’m actually trying to read on the left-hand side.” [/news/Teesside-news/an article] N22n

169 01:36 “As I scroll down there seems to be comments made by... I suppose people who read the news. It’s something which I personally... not detest, but I’ve never done... I see it as pretty pointless. If you’ve got a viewpoint you speak to a person about it I suppose.” [/news/Teesside-news/an article] N22n

172 07:29 “I mean the layout is as standard as before. I still think it’s slightly too busy for me, but, you know, I suppose it needs to generate income somewhere.” [/lifestyle/] N22n

185 03:14 “So they have a picture about the football but not about the raid [a previous article]. Hmm...” [/boro-fc/an article] N25-

187 05:01 “They’ve got questions and answers, that’s good.” [/lifestyle/parenting advice/an article] N25+

Layout

7 00:40 “The one difficulty with the... with this part of the site is that the active story piece is quite small, and quite often, even with a small story like this top one... [/news/Teesside-news/an article] ...it’s split into two [clicks next page], because, presumably, they haven’t got the space to put it all on one, and yet the surrounding material, telling me about First Trans Pennine and various other advertisers has very high graphic values, and so presumably is taking out large amounts of the... story. And yet, this appears to be something peculiar in my experience to the Gazette, because other newspaper sites don’t have this problem. [Goes back to local news] And yet, if you look at the... go back to the local news homepage, which I have now done, it has a great deal more active information, so it seems odd, that you’ve got this limitation on the... on the actual news story.” G3-

8 02:15 “...there is only a couple of dozen lines and a picture... and that it looks as though it probably marks the limit of what it can do in... in one page.” [/news/Teesside-news/an article] G3n

34 04:51 “Okay, the articles, I think, are written in a very... the writing style is not very complicated, which is good to pick up news easily, or what I think. There is a big headline, so you get a... you just read the headline and with pictures so it’s easy to get... easy to get an idea when you just read the headline and

see the picture, and then when you're into the topic... it's not very detailed, but for the purpose of what I want to know I think it's enough." [/news/Teesside-news/an article] N5+

90 02:42 "...so I'll have a look at the second page [an article] which is all text. There is no picture like there was on the previous page. Quite a lot to read. But it's all broken down so it's quite easy to read, it's not just one big paragraph. Which would make me think it would take a long time to read so probably I wouldn't read it." [/news/Teesside-news/an article] N11n

92 03:27 "...and then it just brings up a list of the main stories with the headline written in bold, which makes it quite easy to read, so you know which ones you wanna look at and which ones you don't." [/news/UK-and-world-news/] N11+

95 06:46 "It's quite a long story again, but it's all broken up into paragraphs so it makes easy to read but I probably just skim read it" [/news/Teesside-news/an article] N11+

96 07:12 "There's no pictures on this second page, it's just text [negative tone], so it seems to be a lot more of it than there was last time so I probably won't read that 'cause it's so much. But it is broken up, which does make it easy to read, so... maybe I'll read it later, I'll go back to it." [/news/Teesside-news/an article/p2] N11n

103 02:05 "...but again, I find it there's still a lot of information flashing in front of me, which I don't particularly like. And the writing is really quite small for me anyway." [/north-east-sport/latest-sport-news/an article] N12-

120 01:41 "Okay, so the news story is split over two pages. That's a bit irritating. If there is... there's a lot of space at the bottom for... oh, there's loads of space at the bottom. [...] I think it'd be much better if that was just over one page, if the story was just only one place and you could just read all of it, rather than having to flip to another page." [/news/Teesside-news/an article] N15-

150 06:52 "Again, there's nothing really jumping out for me from the titles of the stories. [...] Maybe the layout... I'd quite like to see some pictures to go with it so it draws your eye to what the story is about." [/news/.../Stockton-focus/] G19-

164 07:27 "...with, say, very short sentences, the longest sentence being maybe two and a half lines. Don't know what's saying about the readers, whether they have got not much attention span. The shortest sentence on here is maybe seven-eight words or so, so keeping it short. But the... [muttering] purposely writing it to dumb down [rushed muttering] really can read things a bit more in-depth and, feeling like they're cheating you with a kind a level of intelligence for you to read more than one sentence at a time without a space." [An article] G20-

171 03:34 "You know there's classifieds and advertisement on the right as usual, as before, but it seems to be pretty clear, you know. Each, each paragraph is set into... a simple sentence. Seems very, very easy to... to follow." [/boro-

fc/an article] N22+

173 07:38 “Also what is very consistent is the Gazette... they use the red and the blue and it does seem to flow throughout the sequence. All the quick links are in red and all the major headlines are in blue and the text seems to be in blue and... So you have, there’s a consistency of that.” [/lifestyle/] N22+

Information architecture

22 12:07 “They’ve got official Web site, facebook, a blog, then I get MySpace or twitter of flickr... Quite a few different connecting to quite popular social network sites, that’s good.” [/news/Middlesbrough-music-live/] G4+

25 00:26 “...and, for example, if I want to find out a specific topic, for example, like elections or what Gordon Brown said this morning I would know where exactly to look at.” [Home] N5+

26 00:38 “The only useful thing, or which I immediately see when I look here is the Google search...” [Home] N5-

37 06:53 “Now this is better. This is how I like to read the news. UK and world news, it’s a good overview with a headline. I think they should do it in... or not they should, but I prefer if news are listed in that way so you can read the headlines and immediately click on it and maybe in these categories like local news and world news... but, the news also don’t seem very... organised according to topics... ” [/news/UK-and-world-news] N5+

51 05:39 “The links all seem to be adverts as well.” [/useful-information/finance/an article] N6-

66 05:30 “So this must be the main menu for all the football news. That’s a good idea, I like having that separate to the sports, so if you’re just interested in Middlesbrough Football Club, things which is what I... I guess if I’d get... if I buy the newspaper I usually check the front headline story, initially, first paragraph, then I flip to the sports, flip to just where I find the Middlesbrough section. So this is interesting to me, I’d go straight to this probably, first. If I knew it was here... Usually, I just follow links of Google or whatever, but I see you can get it off the home page.” [/boro-fc/] G7+

67 06:30 “This is interesting. I guess some of these stories have already appeared on the front page, but it’s good to have it in one place. [/business/] G7+

71 00:15 “Oh, you’ve got a Google search here, that’s a site which I constantly use. I’m glad that they’ve got that on here.” [Home] N8+

80 08:35 “I don’t know how to get to the text of the picture...” [/north-east-history/galleries/a gallery] N9-

91 03:18 “So probably I’ll go back to... I go back to UK and world news, and that’s at the top at one of the tabs which is quite easy to navigate around...” [/news/Teesside-news/ an article] N11+

99 00:30 “There’s very... very little links from there, actually.” [Home; today’s news] N12-

100 00:53 “Looking at the bottom of the homepage to see if there’s anything about

sports pages... Yes, I've found one but it's very, very small and there's an awful lot of links and hyperlinks here. Makes it quite difficult to [incomprehensible]." [Home] N12-

101 01:14 "Just linking on to sports, there's an awful lot of information here. There's a lot of... flash and links at the top... about the army. There's a lot of other links down the side that's quite off-putting, actually." [/north-east-sport/] N12-

104 02:23 "I'm clicking 'back' again to get to the home page... and nothing's happening. [Clicks again] Right, nothing's happening at all. I'll click on the 'home' [clicks]. That takes me back. Okay. That wasn't very good. Normally if I click the 'back' button it will take me back to the home page, but that didn't work. And I find that quite annoying to be honest... [Continues in next comment]" [a search result page] N12-

105 02:44 "...especially when there is so much information on the Web site, on the home page and the various links... I just can't see which... which is which." [Home] N12-

110 07:12 "[reads comments]...but I like the link to these, because there's very little information on there. I've linked to the comment and all I see is the comment, which is exactly what I'd want to see, not lots of different... other links" [/local-news/comments of an entry] N12+

115 06:06 "Some of the links are quite deceiving 'cause it looks like it goes into another area of the Gazette, but then it really goes to another Web site." [Home] N13-

122 02:58 "Ah, so this must be to do with... hmm... communities and things. There's space for blogging and a space for local stories in the middle there. Quite easily accessible as it in the map... and there's a map so you can see where they are all based. That's quite good." [/TS10/] N15+

125 05:02 "And then the sports organised by... different links there down the side ['middle bar']. I think that's a bit of funny place to put it. You kind of... either there's tabs there [points to upper-left corner] or tabs straight down the left. There's the first place I'd look. I think you normally look top left for... or straight across the top for... Oh, they are across the top, yeah [laughs]. Yeah, across the top there, but they're not that... they're quite small, like they'd be better if they were a bit bigger." [/north-east-sport/] N15-

128 06:33 "That's okay, it's quite easy to get to from the... sport pages [/north-east-sport/north-east-cricket/an article] N15+

129 06:40 "Ah, then there's links to pages outside as well [points to middle bar], which is... that's good. It's useful." [/north-east-sport/north-east-cricket/] N15+

138 07:17 "Feels like... It's quite hard to navigate." [/videos-pictures/Teesside-at-night/flickr/] N16-

149 05:38 "With this one there doesn't seem to be any, any linked stories with it [no middle links bar as previously]. At this point I wouldn't mind some information on what had been happening in Afghanistan, get some context of

it, of the related stories that happened. Sometimes, you can get that on some of the sites, which have a... such as the BBC which have a story or one... related news stories come up. I don't seem to find it on here." [/news/uk-and-world-news/an article] G19-

159 03:19 "Obviously it's all chronological, rather than, say, flicks through different aspects [news box], and again more links off the site maybe [points at bottom link box]... give you the same thing more off the site rather than on the site." [/news/Teesside-news/] G20-

160 03:38 "Probably the biggest useful thing, but it's the smallest there at... the toolbar at the top with the different aspects... Environment... Remember when... Campaigns et cetera. That could be a bit more, kind of specific, that's a little bit vague. Tees pride... No idea what's that referring to. And Remember when is just like nostalgic and things like that. It's probably one of the more useful bits but it's probably one of the smaller bits there. As a result probably gets ignored." [Home] G20n

163 05:22 "...and it loads up. Then it goes to a different site. It doesn't even... It's the Gazette site again links you away. Because What's on is probably useful finding out things about the area and things like that, but it takes you away from the Gazette site." [What's on] G20-

178 03:48 "I find the Web site a bit heavy as well. It's too much information going on and that makes me a bit irritated." [/business/] N24-

Diversion

14 00:26 "Not taking too long to load." [/news/Teesside-news/] G4+

18 06:00 "Taking a while to load..." [eveninggazette.newspaperdirect.com/epaper/] G4-

23 12:29 "Doesn't take too long to load." [/news/Middlesbrough-music-live/an article] G4+

27 00:43 "yeah, it's very colourful, therefore... which makes it very confusing to me. Yeah, I've mentioned that I see this site the first time in my life. [Home] N5-

28 01:03 "When I log on to my personal e-mail, there is a news... news button, so always when I log on I get the main five news, and I also have Spiegel online, which is another news site, on this iGoogle, so I can get the five major news of the day and a good overview and this is just too confusing for me [looks at latest news pictures and headlines changing] and I don't know... Well, just when I look at it for some seconds I don't know what, what is the most important thing in the world or in the UK what happened today..." [Home] N5-

33 04:27 "If I'd only open it to see the news, I would just click away. No really, because... so confusing. And there is no overview. [/business/] N5-

40 01:15 "That advert for Lidl keeps distracting me on the right hand side, I keep wanting to look at that." [/news/Teesside-and-tees-valley/local-regeneration/] N6-

43 02:16 "There is far too much information going on at once. Which is good,

‘cause it kind of makes me want to read everything, but it’s taking my mind off what I’m actually looking for.” [Home] N6-

44 02:56 “It’s good that they’ve got... there is not too many distracting colours or anything, it’s just all the images on the right hand side an animations that are doing it.” [An article] N6+

94 05:36 “I’ve just noticed that the date has now changed on these headlines I’m reading. It did change; some of them were on the 14 of July. Which has confused me a little bit, as I thought it was all in chronological... well, going back in a chronological order” [/news/UK-and-world-news/p3] N11-

97 00:06 “My first reaction’s there’s an awful lot of information and I’m really not sure where I’d go first.” [Home] N12-

98 00:25 “However, if I wanted to just read through the news... I’m really quite confused... oh no, there’s today’s news from the Evening Gazette.” [Home] N12-

101 01:14 “Just linking on to sports, there’s an awful lot of information here. There’s a lot of... flash and links at the top... about the army. There’s a lot of other links down the side that’s quite off-putting, actually.” [/north-east-sport/] N12-

105 02:44 “...especially when there is so much information on the Web site, on the home page and the various links... I just can’t see which... which is which.” [Home] N12-

108 04:24 “Oh. Right. There’s a local community Web site with lots of initial post codes and letters and numbers... which really looks quite confusing.” [/gazette-communities/] N12-

109 06:26 “but there’s so much information I’m really finding it quite hard to see if there’s any links anywhere else about market times...” [TS9 articles] N12-

127 06:11 “...but no picture with that [article]. [...] I think it’d be better if it was more like a proper newspaper that... it was, there was pictures and text. There were pictures and text but then, like you didn’t have the adverts. And those moving adverts... really distracting. I don’t like them at all. But this has no pictures at all [the article]. But at least it’s all on one page this time [...] ” [/north-east-sport/north-east-cricket/an article] N15-

139 08:00 “The advertising is really annoying. It’s actually quite big on the screen and so... I’m not used to that... I’m used to two smaller sidebars really, but that’s quite distracting. Oh, I’ll just ignore it.” [/videos-pictures/your-uploaded-pictures/] N16-

141 08:54 “It’s reasonably quick to load up I think, but if I was at home, it might take longer. This is a university computer.” [/videos-pictures/your-uploaded-pictures/] N16+

145 09:33 “It’s taking a while to turn over. A shame. [Long waiting time]” [next page link for headlines] N18-

158 02:57 “Again, I think the advertising is probably distracting more away the site, rather than anything else. Almost as if they want you to click off the site and go

somewhere else" [/news/Teesside-news/] G20-

174 09:15 "I think that the quality of the player is actually quite good actually. It didn't take too long to load up either." [A video] N22+

184 03:08 "Takes a while to load up as well." [Home-article] N25-

186 03:52 "It takes too long to load" [form article to Lifestyle section] N25-

188 05:27 "This takes too long." [From an article to Fashion & Beauty] N25-

Appendix 4.1. Online questionnaire (Study 2)

1 INFORMATION SHEET FOR PARTICIPANTS

Dear participant,

Thank you for taking part in this study. The purpose of the study is to explore how people rate their interaction experience with news sites using a set of questionnaires. You will be asked to interact with a news site of your own choice and, following that, to complete a set of questionnaires about your use of the site. Finally, you will be asked to answer some questions about yourself and your habits regarding Internet use. With your help, I aim to have a better understanding of how to measure interaction experience. It is expected that the results of the research will help in the design of a model of interaction experience. Altogether your participation in the study will take approximately 30 minutes. After browsing a news site of your choice, the completion of the questionnaire will take approximately 10 minutes of your time.

All personal information and data obtained during the study will be treated confidentially. Data will be stored on a university computer in a university building and will only be accessed by the researcher. Data will be kept until the completion of the research.

Your anonymity will be maintained during the research. An identification number will be used to identify your data and no personal information will be published that could identify you as a person.

If you feel uncomfortable with the procedure, you have the right to withdraw at any time before submitting this survey form on the last page, without any personal consequences.

If you have any queries about the purpose and procedure of the study, please feel free to contact the researcher. Once again, thank you for taking part in this study.

Researcher: Gabor Aranyi

School of Social Sciences and Law
Teesside University
TS1 3BA Middlesbrough, UK

E-mail: h8120322@tees.ac.uk

2. INFORMED CONSENT [CHECKBOXES TO EACH STATEMENT]

I have been informed of the purpose of the study as described in the Participant Information Sheet

I have been given the opportunity to ask questions about the study

I have the right to withdraw at any time - without personal consequences - up to the submission of this survey form on the last page

I have been informed that personal information and the data obtained during the study will be treated confidentially. Data will be stored on a university computer in a university building and will only be accessed by the researcher. Data will be kept until the completion of the research

I agree to participate in the study as outlined to me

3. INSTRUCTIONS

I would like to ask you to use a news Web site of your own choice for the next few minutes. You are free to browse any site, as long as it is a Web site with the primary purpose of reporting news (for example: The Guardian, BBC News and The Independent). You can also chose a subject-specific news site if you wish (for example: motoring or music news sites). Only use one news site.

Please do not navigate away from the site of your choice during this session and interact only with the news site. Do not access other Web sites and do not engage in any other activities during this time.

Continue interacting with the news site of your choice for as long as you like

It is absolutely essential that you use a news site of your choice before completing the questionnaire. In order to facilitate the quality of data collected for this survey, please use the site for at least a couple of minutes before proceeding to the questionnaire.

Please do not complete this questionnaire if you do not use news sites at all.

Now, open a new browser window and start using the news site of your choice. Once you are done using the news site and have closed it, please return to this window and click Next to proceed to the questionnaire.

REMEMBER, DO NOT CONTINUE UNTIL YOU HAVE FINISHED USING THE NEWS SITE AND CLOSED IT.

[QUESTIONNAIRES BEGIN]

Please answer the following questions. Remember to indicate an answer in all places.

1. Please give the name of the news site you have been using. [Text box]

2. Please recall your use of the site and indicate the extent to which you have experienced the feelings listed below on a 1 to 7 scale. [PANAS]

During the use of the site I felt...

POS1	Interested	NEG6	Irritable
NEG1	Distressed	POS6	Alert
POS2	Excited	NEG7	Ashamed
NEG2	Upset	POS7	Inspired
POS3	Strong	NEG8	Nervous
NEG3	Guilty	POS8	Determined
NEG4	Scared	POS9	Attentive
NEG5	Hostile	NEG9	Jittery
POS4	Enthusiastic	POS10	Active
POS5	Proud	NEG10	Afraid

Response format: 7-point Likert scale with anchor points ‘not at all’, ‘moderately’ and ‘extremely’.
Item were presented in the order suggested by Watson et al. (1988).

3. Please recall your use of the news site. You will find word pairs below to assist you in the evaluation of the site. Each pair represents extreme contrasts. The possibilities between the extremes enable you to describe the intensity of the quality you choose. Indicate your choice for each word pair by ticking one of the seven circles below the words [ATTRAKDIFF2].

PQ1	Confusing	-	Structured
PQ2	Unpredictable	-	Predictable
PQ3	Impractical	-	Practical
PQ4	Complicated	-	Simple
HQ1	Dull	-	Captivating
HQ2	Tacky	-	Stylish
HQ3	Cheap	-	Premium
HQ4	Unimaginative	-	Creative
BEAUTY	Ugly	-	Beautiful
GOODNESS	Bad	-	Good

Response format: 7-point semantic differential.
Items were presented in the order suggested by Hassenzahl and Monk (2010).

4. Please recall your use of the news site. Indicate to which extent you agree with the following statement [ATTRIBUTION].

My experience in that situation was actually caused by the news site.

Response format: 7-point Likert scale with anchor points “strongly disagree” and “strongly agree”.

5. Please recall your use of the news site. Indicate to which extent you agree with each of the following statements [PERCEIVED ENJOYMENT].

PE1	I find using this news site to be enjoyable.
PE2	The actual process of using this news site is pleasant.
PE3	I have fun using this news site.

Response format: 7-point Likert scale with anchor points ‘strongly disagree’, ‘neutral’ and ‘strongly agree’.

6. During the use of the site... [PERCEIVED DISORIENTATION]

PD1	I felt lost.
PD2	I felt I was going around in circles.
PD3	It was difficult to find a page I had previously viewed.
PD4	Navigating between the pages was a problem.
PD5	I didn't know how to get to my desired location.
PD6	I felt disoriented.
PD7	After browsing for a while I had no idea where to go next.

Response format: 7-point Likert scale with anchor points ‘strongly disagree’ and ‘strongly agree’.

7. Please recall your use of the news site. Indicate to which extent you agree with each of the following statements [PERCEIVED USER-INTERFACE DESIGN].

PUID1	The layout of the site is user-friendly.
PUID2	The layout of the site is in good structure.
PUID3	Overall, the user-interface design of the site is satisfactory.

Response format: 7-point Likert scale with anchor points ‘strongly disagree’, ‘neutral’ and ‘strongly agree’.

8. Please indicate the extent to which you agree or disagree with each of the following descriptions regarding the news site you have used [PERCEIVED AESTHETICS].

CA1	Clean
CA2	Pleasant
CA3	Symmetrical
CA4	Aesthetic
EA1	Original
EA2	Sophisticated
EA3	Spectacular
EA4	Creative

Response format: 7-point Likert scale with anchor points ‘strongly disagree’ and ‘strongly agree’.

9. Please recall your use of the news site. Indicate to which extent you agree with each of the following statements [USEFULNESS OF CONTENT/ADEQUACY OF INFORMATION/ACCESSIBILITY].

UC1	The site provides relevant information.
UC2	The site provides up-to-date information.
UC3	The site provides unique content.
AI1	The site provides comprehensive information.
AI2	The site provides complete content.
AI3	The site provides sufficient information.
AC1	The pages of the site are accessible.
AC2	The pages of the site load quickly.

Response format: 7-point Likert scale with anchor points ‘strongly disagree’, ‘neutral’ and ‘strongly agree’.

10. Please recall your use of the news site and indicate the extent to which you agree or disagree with each of the following statements [BEHAVIOURAL INTENTION TO USE].

BI1	I intend to use the site in the future.
BI2	I predict that I will use the site in the future.

Response format: 7-point Likert scale with anchor points ‘strongly disagree’ and ‘strongly agree’.

11. During my use of the news site I felt... [NEED FULFILMENT]

AUT1	that my choices were based on my true interests and values.
AUT2	free to do things my own way.
AUT3	that my choices expressed my 'true self'.
COMP1	that I was successfully completing difficult tasks and projects.
COMP2	that I was taking on and mastering hard challenges.
COMP3	very capable in what I did.
REL1	a sense of contact with people who care for me and whom I care for.
REL2	close and connected with other people who are important to me.
REL3	a strong sense of intimacy with the people I spent time with.
STIM1	that I was experiencing new sensations and activities.
STIM2	intense pleasure and enjoyment.
STIM3	that I have found new sources and types of stimulation for myself.
POP1	that I was a person whose advice others seek out and follow.
POP2	that I strongly influenced others’ beliefs and behaviour.
POP3	that I had strong impact on what other people did.

Response format: 7-point Likert scale with anchor points ‘strongly disagree’ and ‘strongly agree’.
The presentation order of the items was randomised.

12. How important were the following feelings to you in relation to using the news site [NEED FULFILMENT RELEVANCE]?

RELAUT (Autonomy)	Feeling like you are the cause of your own actions rather than feeling that external forces or pressures are the cause of your actions.
RELCOMP (Competence)	Feeling that you are very capable and effective in your actions rather than feeling incompetent or ineffective.
RELREL (Relatedness)	Feeling that you have regular intimate contact with people who care about you rather than feeling lonely and uncared for.
RELPOP (Popularity)	Feeling that you are liked, respected and have influence over others rather than feeling like a person whose advice and opinions nobody is interested in.
RELSTIM (Stimulation)	Feeling that you get plenty of enjoyment and pleasure rather than feeling bored and understimulated.

Response format: 7-point Likert scale with anchor points ‘not important at all’ and ‘extremely important’.

[DEMOGRAPHICS, INTERNET-USE BEHAVIOUR AND NEWS-SITE USE BEHAVIOUR QUESTIONS]

1. How old are you? [Text box]

2. What is your gender? [Radio buttons]

Male
Female

3. How long have you been using the Internet (years)? [Text box]

4. How often do you log on to the Internet? (Select one answer) [Radio buttons]

Weekly
2-3 times a week
4-6 times a week
Daily
2-3 times a day
More than 3 times a day

5. Once on line, how much time do you spend on average per visit to the Internet? (Select one answer) [Radio buttons]

1-10 minutes
About 15 minutes
About 30 minutes
About 45 minutes
About 1 hour
Several hours

6. How often do you visit online news sites? (Select one answer) [Radio buttons]

Weekly
2-3 times a week
4-6 times a week
Daily
2-3 times a day
More than 3 times a day

7. How often do you visit the news site you have been using during this session? (Select one answer) [Radio buttons]

Weekly
2-3 times a week
4-6 times a week
Daily
2-3 times a day
More than 3 times a day

8. How much time do you spend on online news sites per visit? (Select one answer) [Radio buttons]

1-4 minutes
About 5 minutes
About 10 minutes
About 15 minutes
About 30 minutes
More than 30 minutes

9. What kind of devices do you use to access news sites? (Select all that apply) [Checkboxes]

Desktop Computer

Laptop Computer
Mobile Phone
PDA
Other (please specify) [Text box]

[QUESTIONS END]

Thank you for completing this survey.

Please provide your e-mail address in the text box below to enter the prize draw. If you are a psychology student, provide your student e-mail address so it can be used to give you SONA credit for participating in this research. If you wish to withdraw from the study after the completion of this survey, your e-mail address will be used to identify your data. If you do not leave a valid e-mail address, you will not enter the prize draw and you will not be allowed to withdraw from the study after submitting this form.

[Text box]

Appendix 4.2. Cross-loadings of items of measures (Study 2)

Item	Scale																
	AC	AES	AI	ATT	BUSE	BEAU	BI	GOOD	HQ	NEG	PD	PE	POS	PQ	PUID	UC	USE
AC1	0.92	0.40	0.50	0.13	0.12	0.17	0.35	0.29	0.25	-0.12	-0.45	0.31	0.19	0.46	0.55	0.55	0.14
AC2	0.92	0.34	0.40	0.04	0.13	0.14	0.38	0.31	0.23	-0.11	-0.43	0.33	0.17	0.43	0.45	0.45	0.14
AI1	0.50	0.52	0.92	0.06	0.08	0.21	0.36	0.30	0.35	-0.08	-0.26	0.34	0.34	0.36	0.42	0.60	0.12
AI2	0.41	0.50	0.92	0.06	0.10	0.21	0.35	0.36	0.40	-0.17	-0.20	0.34	0.24	0.33	0.36	0.56	0.13
ATT	0.09	0.15	0.06	1.00	0.14	0.19	0.07	0.15	0.15	0.13	-0.01	0.19	0.21	0.09	0.13	0.03	0.13
BEAU	0.17	0.54	0.23	0.19	0.13	1.00	0.24	0.54	0.61	-0.03	-0.11	0.41	0.34	0.32	0.36	0.14	0.17
BI1	0.39	0.42	0.38	0.05	0.31	0.25	0.97	0.40	0.35	-0.11	-0.39	0.49	0.28	0.42	0.54	0.46	0.36
BI2	0.37	0.37	0.37	0.09	0.28	0.22	0.97	0.41	0.36	-0.19	-0.39	0.47	0.24	0.41	0.48	0.44	0.28
CA1	0.49	0.64	0.52	0.06	0.21	0.22	0.37	0.34	0.32	-0.09	-0.31	0.32	0.24	0.41	0.47	0.48	0.20
CA2	0.46	0.83	0.51	0.15	0.16	0.47	0.44	0.55	0.57	-0.10	-0.29	0.60	0.41	0.41	0.60	0.45	0.23
CA3	0.18	0.58	0.26	0.12	0.12	0.27	0.17	0.21	0.27	0.07	-0.03	0.26	0.26	0.17	0.27	0.19	0.18
CA4	0.35	0.85	0.43	0.13	0.21	0.51	0.38	0.43	0.56	-0.07	-0.22	0.46	0.32	0.34	0.53	0.30	0.25
EA1	0.24	0.80	0.42	0.12	0.18	0.40	0.27	0.31	0.45	0.03	-0.10	0.37	0.26	0.22	0.39	0.27	0.20
EA2	0.29	0.72	0.45	0.05	0.03	0.29	0.33	0.35	0.47	-0.03	-0.19	0.34	0.25	0.25	0.42	0.39	0.03
EA3	0.15	0.78	0.36	0.18	0.14	0.53	0.18	0.40	0.56	0.10	-0.02	0.39	0.38	0.20	0.36	0.16	0.24
EA4	0.28	0.85	0.42	0.06	0.18	0.49	0.30	0.38	0.58	-0.03	-0.19	0.46	0.32	0.30	0.51	0.31	0.22
GOOD	0.32	0.50	0.36	0.15	0.24	0.54	0.42	1.00	0.62	-0.18	-0.35	0.53	0.30	0.52	0.47	0.34	0.27
HQ1	0.26	0.50	0.36	0.14	0.13	0.42	0.30	0.55	0.80	-0.01	-0.26	0.47	0.31	0.43	0.38	0.30	0.20
HQ2	0.17	0.47	0.32	0.12	0.08	0.51	0.28	0.42	0.80	-0.01	-0.16	0.40	0.28	0.36	0.36	0.24	0.09
HQ3	0.18	0.38	0.25	0.14	0.12	0.33	0.30	0.45	0.69	-0.08	-0.19	0.29	0.18	0.27	0.25	0.25	0.08
HQ4	0.20	0.60	0.33	0.09	0.13	0.62	0.28	0.49	0.81	0.02	-0.18	0.45	0.34	0.33	0.40	0.22	0.16
NEG1	-0.01	-0.01	-0.04	0.11	-0.11	0.01	-0.10	-0.14	0.04	0.66	0.15	-0.16	0.01	-0.11	-0.06	-0.08	-0.11
NEG10	-0.03	0.01	-0.07	0.13	0.01	-0.01	-0.02	-0.12	0.06	0.70	0.10	-0.02	0.10	-0.06	0.02	-0.12	0.03
NEG2	0.08	-0.03	0.03	0.04	-0.08	0.01	-0.01	-0.01	0.10	0.56	0.02	-0.08	0.09	-0.03	0.04	-0.03	-0.01
NEG3	-0.02	-0.03	-0.03	0.06	0.09	-0.08	-0.07	-0.15	-0.02	0.67	0.24	-0.07	0.07	-0.10	-0.10	-0.09	0.10
NEG4	0.00	0.08	0.00	0.12	0.04	0.00	-0.03	-0.05	0.03	0.67	0.09	0.03	0.16	-0.10	0.05	-0.06	0.07
NEG5	-0.05	-0.01	-0.11	0.15	0.09	0.03	-0.03	-0.07	0.03	0.66	0.14	-0.04	0.17	-0.11	0.00	-0.12	0.11
NEG6	-0.12	-0.08	-0.21	0.08	0.06	-0.06	-0.06	-0.15	-0.05	0.69	0.19	-0.05	0.09	-0.16	-0.06	-0.13	0.04
NEG7	-0.09	-0.04	-0.14	0.00	0.04	-0.04	-0.14	-0.13	-0.02	0.74	0.21	-0.08	0.12	-0.22	-0.10	-0.21	0.09
NEG8	-0.07	-0.02	-0.02	0.11	-0.06	0.02	-0.07	-0.16	-0.08	0.70	0.26	-0.10	0.15	-0.14	-0.10	-0.10	-0.01
NEG9	-0.20	0.00	-0.11	0.10	0.01	-0.01	-0.26	-0.12	-0.03	0.72	0.30	-0.06	0.19	-0.21	-0.17	-0.20	0.05

Appendix 4.2

Item	Scale																
	AC	AES	AI	ATT	BUSE	BEAU	BI	GOOD	HQ	NEG	PD	PE	POS	PQ	PUID	UC	USE
BUSE	0.14	0.20	0.10	0.14	**1.00**	0.13	0.30	0.24	0.15	0.03	-0.16	0.37	0.26	0.23	0.19	0.09	0.81
PD1	-0.41	-0.24	-0.23	-0.04	-0.20	-0.18	-0.39	-0.42	-0.29	0.28	0.88	-0.41	-0.19	-0.58	-0.53	-0.27	-0.19
PD2	-0.38	-0.23	-0.29	-0.05	-0.14	-0.13	-0.35	-0.40	-0.30	0.28	0.85	-0.37	-0.19	-0.48	-0.43	-0.32	-0.12
PD3	-0.35	-0.09	-0.21	0.08	-0.03	0.02	-0.24	-0.11	-0.10	0.16	0.73	-0.12	-0.04	-0.25	-0.31	-0.19	-0.03
PD4	-0.43	-0.14	-0.16	-0.05	-0.11	0.00	-0.28	-0.18	-0.15	0.19	0.81	-0.18	-0.07	-0.35	-0.34	-0.23	-0.05
PD5	-0.38	-0.17	-0.19	-0.03	-0.11	-0.05	-0.29	-0.25	-0.16	0.25	0.85	-0.25	-0.09	-0.40	-0.44	-0.19	-0.11
PD6	-0.46	-0.18	-0.17	0.02	-0.12	-0.10	-0.42	-0.25	-0.19	0.31	0.89	-0.26	-0.11	-0.43	-0.46	-0.26	-0.10
PD7	-0.39	-0.23	-0.17	0.03	-0.16	-0.08	-0.32	-0.25	-0.20	0.18	0.80	-0.32	-0.14	-0.36	-0.42	-0.20	-0.20
PE1	0.29	0.47	0.34	0.15	0.34	0.33	0.52	0.48	0.46	-0.15	-0.33	0.92	0.45	0.42	0.51	0.37	0.38
PE2	0.38	0.51	0.34	0.22	0.33	0.39	0.42	0.50	0.50	-0.04	-0.38	0.92	0.48	0.43	0.55	0.36	0.35
PE3	0.27	0.50	0.33	0.15	0.32	0.39	0.40	0.46	0.47	-0.07	-0.24	0.88	0.53	0.30	0.46	0.29	0.35
POS1	0.30	0.35	0.35	0.07	0.26	0.21	0.49	0.35	0.38	-0.09	-0.33	0.51	0.64	0.39	0.37	0.38	0.27
POS10	0.11	0.29	0.19	0.19	0.14	0.27	0.05	0.14	0.19	0.19	0.01	0.32	0.70	0.08	0.18	0.10	0.17
POS2	0.10	0.30	0.19	0.13	0.25	0.32	0.18	0.19	0.26	0.16	-0.06	0.45	0.68	0.11	0.17	0.09	0.26
POS3	0.10	0.25	0.15	0.14	0.07	0.24	0.07	0.12	0.15	0.23	-0.05	0.21	0.58	0.05	0.09	0.05	0.14
POS4	0.06	0.24	0.19	0.15	0.21	0.23	0.12	0.18	0.22	0.09	-0.08	0.41	0.71	0.14	0.15	0.16	0.20
POS5	0.03	0.27	0.14	0.16	0.13	0.28	0.07	0.21	0.23	0.10	-0.02	0.34	0.68	0.13	0.13	0.09	0.15
POS6	0.16	0.18	0.15	0.15	0.09	0.10	0.09	0.12	0.20	0.21	-0.12	0.24	0.59	0.16	0.23	0.14	0.02
POS7	0.10	0.32	0.27	0.15	0.23	0.26	0.20	0.26	0.28	0.10	-0.07	0.44	0.78	0.17	0.21	0.20	0.22
POS8	0.04	0.26	0.15	0.17	0.15	0.27	0.03	0.12	0.17	0.30	0.01	0.22	0.71	0.09	0.12	0.11	0.14
POS9	0.15	0.23	0.20	0.17	0.12	0.20	0.17	0.16	0.22	0.13	-0.13	0.29	0.68	0.13	0.25	0.20	0.12
PQ1	0.43	0.45	0.37	0.10	0.20	0.29	0.36	0.51	0.38	-0.16	-0.48	0.46	0.29	0.85	0.60	0.36	0.22
PQ2	0.07	0.01	0.01	-0.02	0.09	0.12	0.12	0.11	0.15	-0.12	-0.10	0.04	0.04	**0.39**	0.07	0.11	0.05
PQ3	0.38	0.26	0.31	0.07	0.19	0.18	0.41	0.38	0.34	-0.22	-0.38	0.33	0.21	0.81	0.43	0.40	0.14
PQ4	0.38	0.23	0.23	0.06	0.16	0.32	0.27	0.39	0.39	-0.13	-0.40	0.24	0.09	0.76	0.44	0.25	0.19
PUID1	0.52	0.50	0.38	0.15	0.16	0.31	0.48	0.44	0.41	-0.10	-0.47	0.51	0.29	0.59	0.92	0.43	0.21
PUID2	0.49	0.58	0.42	0.12	0.20	0.33	0.51	0.43	0.42	-0.09	-0.49	0.55	0.30	0.55	0.95	0.42	0.25
PUID3	0.51	0.59	0.40	0.10	0.16	0.37	0.50	0.44	0.44	-0.12	-0.49	0.51	0.28	0.55	0.93	0.46	0.17
USE	0.15	0.26	0.13	0.13	0.81	0.17	0.33	0.27	0.18	0.06	-0.15	0.40	0.27	0.22	0.22	0.10	1.00
UC1	0.51	0.42	0.59	0.10	0.12	0.15	0.42	0.39	0.30	-0.17	-0.30	0.39	0.28	0.40	0.50	0.93	0.11
UC2	0.49	0.34	0.57	-0.05	0.02	0.09	0.43	0.23	0.29	-0.20	-0.24	0.29	0.19	0.35	0.34	0.90	0.07

Note. Grey background indicates the loadings of items belonging to a particular scale. Bold numbers indicate loadings smaller than .50. AC: accessibility. AES: aesthetics. AI: adequacy of information. ATT: attribution. BUSE: baseline use frequency. BEAU:

beauty. BI: behavioural intention. GOOD: goodness. HQ: hedonic quality. NEG: negative affect. PD: perceived disorientation. PE: perceived enjoyment. POS: positive affect. PQ: pragmatic quality. PUID: perceived user-interface design. UC: usefulness of content. USE: use frequency.

Appendix 6.1. Procedure of the experimental program and scales (Study 3)

[PROGRAM STARTED BY RESEARCHER]

Participant types in participant identifier

Participant selects experimental condition (Test Version 1 or 2)

Note. Participant identifier and experimental condition are provided on the participant information sheet (see Appendix 6.2) by the researcher.

Participant clicks ‘proceed’

[Experiment instructions appear on the screen]

Experiment Instructions

Thank you for your participation. This experiment aims to test people’s experiences while using news Web sites. First, you will be asked to complete a computerised task and answer some questions. Following this, you will be instructed to use a news sites for 10 minutes. After your use of the site, you will be asked to answer a series of questions about your experience. These questions are not to test your abilities, but the quality of the site. Finally, you will be asked to answer some questions about demographics and your Internet-use habits.

All instructions will be presented on the computer screen before each task. Please work alone and do not talk to others during the experiment. If you encounter any problems during the session, indicate your need for help to the researcher by raising your hand. Once you are finished, please remain at your workstation to allow other participants to finish undisturbed. Please switch off your mobile phone now. You will not be able to make phone calls during the experiment. The experiment takes about 30 minutes in total to complete.

Participant clicks ‘proceed’

Appendix 6.1

[CENTRALITY OF VISUAL PRODUCT AESTHETICS]

Instructions:

You will find statements below that describe people's attitude towards visual product design. Please indicate the extent to which you find each of the statements truthful to yourself. If the statement is extremely uncharacteristic of you (not at all truthful) please choose 1 on the left; if the statement is extremely characteristic of you (very truthful) then please choose 7 on the right. Of course, a statement may be neither extremely uncharacteristic nor extremely characteristic of you; if so, please use the number in the middle of the scale that describes the best fit. Please keep the following scale in mind as you rate each of the statements below: 1 = extremely uncharacteristic; 4 = somewhat characteristic; 7 = extremely characteristic.

Scale items:

CVPA1	Owning products that have superior designs makes me feel good about myself.
CVPA 2	I enjoy seeing displays of products that have superior designs.
CVPA 3	A product's design is a source of pleasure for me.
CVPA 4	Beautiful product designs make our world a better place.
CVPA 5	Being able to see subtle differences in product designs is one skill that I have developed over time.
CVPA 6	I see things in a product's design that other people tend to pass over.
CVPA 7	I have the ability to imagine how a product will fit in with designs of other things I already own.
CVPA 8	I have a pretty good idea of what makes one product look better than its competitors.
CVPA 9	Sometimes the way a product looks seems to reach out and grab me.
CVPA 10	If a product's design really 'speaks' to me I feel that I must buy it.
CVPA 11	When I see a product that has a really great design I feel a strong urge to buy it.

Note. Value 1-4, acumen 5-8 and response 9-11. One item is presented at a time. Instructions are present above each item.

[MENTAL-ROTATION TEST]

Instructions:

This is to test your ability to look at a drawing of a given object and find the same object within a set of dissimilar objects. The only difference between the original object and the chosen object will be that they are presented at different angles. An illustration of this principle is given below where the same single object is given in five different positions. Look at each of them to satisfy yourself that they are only presented at different angles from one another.

[Five pictures of the same object are presented at different angles]

Below are two drawings of new objects. They cannot be made to match the above five drawings. Satisfy yourself that they are different from the above.

[Two pictures of the same object in different positions, different from the object above, are presented]

Now let's do some sample problems. For each problem there is a primary object in the far left. You are to determine which two of four objects to the right are the same object given on the far left. In each problem always two of the four drawings are the same object as the one on the left. You are to put click on the boxes below the correct ones and not click on the incorrect ones. The first sample problem is done for you.

[Completed sample picture presented]

Participant clicks 'next' for the next screen

Instructions:

Do the rest of the sample problems yourself. Which two drawings of the four on the right show the same object as the one on the left? There are always two and only two correct answers for each problem. Click under the two correct drawings.
This is problem 1 out of 3.
Please click on 'next' to move to the next page

The three practice problems are presented separately on three pages. Instructions are repeated above each problem. After the three practice problems, a screen is presented with the correct answers and additional instructions.

Additional instructions:

You will have 3 minutes for the main test, which follows now. Remember: there are always two and only two correct answers for each item. Work as quickly as you can without sacrificing accuracy. Your score on this test will reflect both the correct and incorrect responses. Therefore, it will not be to your advantage to guess unless you have some idea which choice is correct.

Participant clicks 'next'

Ten trial problems are presented after one another. After three minutes, the experimental program automatically redirects to the next phase. The following instructions are repeated above each item:

Which two drawings of the four on the right show the same object as the one on the left? There are always two and only two correct answers for each problem. Click under the two correct drawings.
This is problem 1 out of 10.
Please click on 'next' to move to the next page.

[SCREENSHOTS AND ATTRAKDIFF2 BEFORE USE]

Announcement screen:

Five screenshots of the web site follow. Please go through them briefly and answer the questions that follow.

Participant clicks 'proceed' for the first screenshot

Five screenshots are presented on separate pages (screenshots of BBC for condition 1 and NZH for condition 2). The following instructions are presented above the picture on each screen:

Please note the following page (1/5) of the web site; you do not need to memorise this page. Please click on 'next' to move to the next page.

Announcement screen:

Please answer the questions about the web site on the following screens.

Participant clicks 'proceed' for the first item of the AttrakDiff2 abridged version

Instructions:

You will find word pairs below to assist you in the evaluation of the site. Each pair represents extreme contrasts. The possibilities between the extremes enable you to describe the intensity of the quality you choose. Indicate your choice for each word pair by pressing one of the seven buttons below the words.
Click on the button that most closely corresponds to your opinion.

I judge the web site to be…

Items:

PQ1	Confusing	-	Structured
PQ2	Unpredictable	-	Predictable
PQ3	Impractical	-	Practical
PQ4	Complicated	-	Simple
HQ1	Dull	-	Captivating
HQ2	Tacky	-	Stylish
HQ3	Cheap	-	Premium
HQ4	Unimaginative	-	Creative
BEAUTY	Ugly	-	Beautiful
GOODNESS	Bad	-	Good

Note. Response format: 7-point semantic differential. One item is presented at a time. Instructions are present above each item.

[INSTRUCTIONS FOR NAVIGATING THE WEB SITE IN THE EXPERIMENT]

Announcement screen:

On the following pages some tips will be presented about navigating the web site that is used in the experiment.

Participant clicks 'proceed'

Three screens follow with screenshots of a page of the Teesside University web site including navigation tips. The following instructions are presented above each screen:

Here are some practical tips (1/3) to navigate the web site. You do not need to memorise this layout. Please click 'next' to move to the next page.

Verbal instructions on the screenshots:

There is no back button in the experiment. In order to go back, you need to click with the right mouse button and then select the option 'back' (see below).

There is no forward button in the experiment. In order to go forward, you need to click with the right mouse button and then select the option 'forward' (see below).

In order to refresh the page you need to click the right mouse button and then select the option 'refresh' (see below).

Appendix 6.1

Instruction screen appears

Instructions for condition 1 (BBC):

Now it is time to use the news site.
Use the BBC site (http://www.bbc.co.uk/) to find out what's on.
The experimental software will automatically redirect to questions about your experience after 10 minutes.
Click 'show news site' to proceed.

Instructions for condition 2 (NZH):

Now it is time to use the news site.
Imagine that you are going to spend a year of studying or working in New Zealand.
Use the New Zealand Herald (http://www.nzherald.co.nz/) to find out what's on.
The experimental software will automatically redirect to questions about your experience after 10 minutes.
Click 'show news site' to proceed.

Participant clicks 'show web site'.

[BROWSING THE NEWS SITE]

Ten minutes of free browsing follows. The news site (BBC or NZH) is presented in full screen with no browser controls present. The experimental software automatically redirects to questions after ten minutes.

Announcement screen:

This message concludes your use of the web site. Please answer the questions on the following pages.

Participant clicks 'proceed'

[POSITIVE AND NEGATIVE AFFECT SCHEDULE]

Instructions

Please recall your use of the site and indicate the extent to which you have experienced the feelings listed below on a 1 to 7 scale. If you have not experienced a feeling at all during your use of the site, please choose 1 on the left. If you have experienced a feeling to a large extent, please choose 7 on the right. If neither is the case, choose a number between the two extremes to scale your response.
Click the button that most closely corresponds to your opinion.

During the use of the site I felt...

POS1	Interested	NEG6	Irritable
NEG1	Distressed	POS6	Alert
POS2	Excited	NEG7	Ashamed
NEG2	Upset	POS7	Inspired
POS3	Strong	NEG8	Nervous
NEG3	Guilty	POS8	Determined
NEG4	Scared	POS9	Attentive
NEG5	Hostile	NEG9	Jittery
POS4	Enthusiastic	POS10	Active
POS5	Proud	NEG10	Afraid

Note. One item is presented at a time. Instructions are present above each item. Items are presented mixed.

[ATTRAKDIFF2 AFTER USE]

The same applies as to AttrakDiff2 before use (see above).

[PERCEIVED ENJOYMENT]

Instructions:

Please recall your use of the site. Indicate to which extent you agree with each of the following statements on a 1 to 7 scale.
Click on the button that most closely corresponds to your opinion.

Items:	
PE1	I find using this news site to be enjoyable.
PE2	The actual process of using this news site is pleasant.
PE3	I have fun using this news site.

Note. One item is presented at a time. Instructions are present above each item.

[PERCEIVED DISORIENTATION]

Instructions:

Please answer the following questions about your experience of using the web site. Click on the button that most closely corresponds to your opinion.

Items:	
PD1	I felt lost.
PD2	I felt I was going around in circles.
PD3	It was difficult to find a page I had previously viewed.
PD4	Navigating between the pages was a problem.
PD5	I didn't know how to get to my desired location.
PD6	I felt disoriented.
PD7	After browsing for a while I had no idea where to go next.

Note. Anchor points: ‘never’ to the left and ‘always’ to the right. One item is presented at a time. Instructions are present above each item.

[PERCEIVED AESTHETICS]

Instructions:

Please indicate the extent to which you agree or disagree with each of the following descriptions regarding the visual design of the site.
Click on the button that most closely corresponds to your opinion.

Items:	
CA1	Clean
CA2	Pleasant
CA3	Symmetrical
CA4	Aesthetic
EA1	Original
EA2	Sophisticated
EA3	Spectacular
EA4	Creative

Note. Response format: 7-point Likert scale with anchor points ‘strongly disagree’ and ‘strongly agree’. One item is presented at a time. Instructions are present above each item.

[USEFULNESS OF CONTENT, ADEQUACY OF INFORMATION AND TRUST]

Instructions:

Please recall your use of the news site. Indicate to which extent you agree with each of the following statements. Please keep the following scale in mind as you rate each of the statements below: 1 = strongly disagree; 4 = neutral; 7 = strongly agree. Click on the button that most closely corresponds to your opinion.

Items:	
UC1	The site provides relevant information.
UC2	The site provides up-to-date information.
UC3	The site provides unique content.
AI1	The site provides comprehensive information.
AI2	The site provides complete content.
AI3	The site provides sufficient information.
TRU1	I believe the site provides truthful information.
TRU2	I trust the information presented on the site.
TRU3	The site provides reliable information.
TRU4	I trust the competence of journalists working for the site.

Note. One item is presented at a time. Instructions are present above each item.

Appendix 6.1

[BEHAVIOURAL INTENTION TO USE]

Instructions for condition 1 (BBC):

Please recall your use of the news site and indicate the extent of your agreement with each of the following statements. Please keep the following scale in mind as you rate each of the statements below: 1 = strongly disagree; 7 = strongly agree.

Instructions for condition 2 (NZH):

Imagine that you are going to live in New Zealand. Indicate the extent of your agreement with each of the following statements accordingly. Please keep the following scale in mind as you rate each of the statements below: 1 = strongly disagree; 7 = strongly agree.

Items (for both conditions):

BI1	I intend to use the site in the future.
BI2	I predict that I will use the site in the future.

Note. One item is presented at a time. Instructions are present above each item.

[SATISFACTION]

Instructions:

Please answer the following questions about the news site you have been using. Click on the button that most closely corresponds to your opinion.

Items:

SAT1	Considering all your experience with the site, how satisfied are you in general?
SAT2	To what degree do you consider that the site fulfils your expectations?
SAT3	Imagine a news site that is perfect in all aspects. How close to this ideal do you consider this site to be?

Note. Items are measured on a 1-7 scale. Anchor points: SAT1 completely dissatisfied – completely satisfied; SAT2 much less than expected – much more than expected; SAT3 very far away – very close. One item is presented at a time. Instructions are present above each item.

[DEMOGRAPHICS, INTERNET USE BEHAVIOUR AND NEWS SITE USE BEHAVIOUR]

Each question is presented on a separate page.

1. Are you male or female? (Please select one answer.)

2. How often do you visit news sites? (Please select one answer.)

Weekly
2-3 times a week
4-6 times a week
Daily
2-3 times a day
More than 3 times a day

3. What is you highest level of education? (Please select one answer.)

Primary school
GCSE or equivalent
Apprenticeship
NVQ Level 1 or 2
Two or more A-levels
NVQ Level 3, 4 or 5
Degree (for example BA or BSc)
Higher degree (for example PGCE, MA, MSc or PhD)

4. In which year did you start to use the Internet? (Please type a number.) [Only numbers between 1990 and 2011 are accepted.]

5. Approximately how many hours do you use the Internet per week? (Please type a number.) [Only numbers between 0 and 168 are accepted.]

6. Approximately how many minutes do you spend on news sites per visit? (Pease type a number.) [Only numbers between 0 and 1440 are accepted.]

7. Approximately how many minutes do you spend on news sites per day? (Pease type a number.) [Only numbers between 0 and 1440 are accepted.]

8. What is your age? (Please type a number.) [Only numbers between 18 and 80 are accepted.]

Appendix 6.1

Announcement screen:

This is the end of the experiment. Thank you for your participation!

Participant clicks ‘click here to finish’ to close program

Appendix 6.2. Participant information sheet and consent form (Study 3)

Participant information sheet

Testing a psychological model of end-users' interaction with news Web sites.

Teesside University
School of Social Sciences and Law

Participant identifier: .....

Experimental condition: .....

Dear Participant,

Thank you for your participation. This experiment aims to test people's experiences while using news Web sites. You will be asked to complete a short task on the computer screen and answer some questions. Following this, you will be instructed to use a news site and answer a series of questions regarding your experience, Internet-use habits and demographic details. It is expected that the results of the research will help in the design of better news sites. The procedure is expected to take about 30 minutes to complete.

No risk or discomfort is expected as a result of your participation. Your anonymity will be maintained in the research process. An identification number rather than your name will be used to identify your data and no personal information will be published. If you feel uncomfortable with the procedure, you have the right to withdraw at any time during the research without any personal consequences, up to the start of data analysis. Personal information and the data obtained during the experiment will be stored on a password-protected university server and treated confidentially. Data will only be accessed by the researcher and will be kept until the completion of the research. If you have any queries about the purpose and procedure of the experiment, please feel free to ask the researcher. Once again, thank you for taking part in this study.

Researcher: Gabor Aranyi

Contact e-mail: G.Aranyi@tees.ac.uk

Appendix 6.2

Informed Consent

Testing a psychological model of end-users' interaction with news Web sites.

I confirm, that (please tick)
I have been informed of the purpose of the study.
I have been given the opportunity to ask questions about the study.
I have the right to withdraw at any time, without personal consequences, up to the start of data analysis.
Personal information and the data obtained during the experiment will be stored on a password-protected university server and treated confidentially. Data will only be accessed by the researchers. Data will be kept until the completion of the research.
I agree to participate in the study as outlined to me.

.....	.....	.....
Your name	Signature	Date
.....	.....	.....
Researcher	Signature	Date

✂.....

Informed Consent

Testing a psychological model of end-users' interaction with news Web sites.

I confirm, that (please tick)

I have been informed of the purpose of the study.

I have been given the opportunity to ask questions about the study.

I have the right to withdraw at any time, without personal consequences, up to the start of data analysis.

Personal information and the data obtained during the experiment will be stored on a password-protected university server and treated confidentially. Data will only be accessed by the researchers. Data will be kept until the completion of the research.

I agree to participate in the study as outlined to me.

....

Your name Signature Date

....

Researcher Signature Date

Appendix 6.3. Loadings and cross-loadings for the measurement model (Study3)

Item	Scale																					
	ACU	AES	AI	B1	B2	BI	G1	G2	HQ1	HQ2	NEG	PD	PE	POS	PQ1	PQ2	RES	SA	SAT	TRU	UC	VAL
AI1	0.09	0.47	0.90	0.25	0.27	0.29	0.49	0.39	0.47	0.44	-0.08	-0.27	0.43	0.43	0.47	0.37	0.23	0.01	0.55	0.49	0.60	0.23
AI2	-0.01	0.60	0.90	0.33	0.33	0.35	0.52	0.46	0.46	0.47	-0.17	-0.41	0.44	0.41	0.44	0.50	0.22	-0.05	0.63	0.49	0.49	0.11
AI3	-0.01	0.51	0.89	0.28	0.31	0.30	0.41	0.47	0.42	0.45	-0.15	-0.33	0.43	0.39	0.44	0.29	0.22	0.05	0.54	0.43	0.52	0.19
B1	0.17	0.63	0.32	1.00	0.60	0.38	0.68	0.54	0.75	0.63	-0.28	-0.41	0.53	0.45	0.40	0.48	0.01	0.05	0.52	0.16	0.35	0.04
B2	0.16	0.68	0.34	0.60	1.00	0.47	0.54	0.68	0.50	0.73	-0.33	-0.49	0.62	0.64	0.25	0.57	-0.01	0.05	0.57	0.05	0.37	0.05
BI1	0.22	0.53	0.36	0.38	0.48	0.99	0.51	0.69	0.44	0.55	-0.30	-0.39	0.54	0.49	0.29	0.46	0.14	0.18	0.66	0.33	0.56	0.03
BI2	0.28	0.49	0.34	0.36	0.46	0.99	0.52	0.68	0.44	0.53	-0.31	-0.38	0.52	0.45	0.32	0.43	0.14	0.21	0.66	0.34	0.53	0.07
CA1	-0.03	0.69	0.39	0.40	0.36	0.39	0.58	0.53	0.37	0.51	-0.18	-0.30	0.42	0.36	0.33	0.48	-0.05	0.02	0.55	0.15	0.38	-0.04
CA2	0.01	0.83	0.48	0.52	0.52	0.58	0.67	0.79	0.54	0.75	-0.31	-0.54	0.69	0.59	0.43	0.66	0.07	0.07	0.76	0.14	0.60	-0.07
CA3	-0.02	**0.47**	0.23	0.23	0.21	0.17	0.31	0.17	0.22	0.28	0.00	-0.17	0.08	0.21	0.21	0.30	0.11	-0.11	0.30	0.14	0.34	0.14
CA4	-0.03	0.79	0.54	0.48	0.50	0.40	0.48	0.61	0.43	0.69	-0.28	-0.47	0.48	0.38	0.38	0.54	0.11	0.07	0.54	0.15	0.41	0.07
CVPA01	0.37	0.08	0.16	0.01	0.03	0.04	0.20	-0.05	0.15	0.05	0.13	0.12	0.00	0.16	0.06	-0.05	0.30	-0.01	0.07	0.16	0.16	0.79
CVPA02	0.43	0.06	0.15	0.03	0.09	0.00	0.16	-0.12	0.13	-0.02	0.21	0.06	0.02	0.24	0.08	-0.05	0.45	-0.02	0.10	0.17	0.13	0.86
CVPA03	0.51	0.13	0.17	0.07	0.01	0.08	0.18	-0.03	0.23	0.09	0.15	0.14	-0.01	0.17	-0.01	-0.08	0.48	0.03	0.10	0.25	0.29	0.90
CVPA04	0.51	0.03	0.17	0.02	0.05	0.04	0.10	-0.04	0.18	0.03	0.23	-0.02	0.03	0.17	-0.01	-0.01	0.44	0.04	0.11	0.26	0.16	0.79
CVPA05	0.88	0.09	0.04	0.28	0.21	0.25	0.24	0.18	0.31	0.14	-0.09	-0.08	0.28	0.29	0.13	0.12	0.28	0.20	0.25	0.09	0.24	0.47
CVPA06	0.86	0.10	0.03	0.09	0.18	0.23	0.18	0.10	0.14	0.10	0.07	-0.05	0.16	0.27	0.19	0.10	0.29	0.12	0.15	0.15	0.12	0.44
CVPA07	0.54	-0.23	-0.17	-0.26	-0.17	0.03	-0.11	-0.14	-0.12	-0.13	0.20	0.20	-0.05	0.04	-0.16	-0.22	0.36	-0.03	-0.05	0.08	-0.03	0.35
CVPA08	0.51	-0.06	0.02	-0.04	-0.12	0.03	-0.02	-0.12	-0.05	-0.12	0.09	0.11	-0.16	0.02	0.00	-0.14	0.44	-0.07	-0.01	0.32	0.04	0.36
CVPA09	0.42	0.09	0.20	-0.07	-0.08	0.06	0.02	0.03	0.08	0.05	0.07	0.11	0.02	0.15	0.07	-0.05	0.72	0.09	0.13	0.23	0.32	0.43
CVPA10	0.28	0.18	0.22	0.09	0.08	0.15	0.18	0.14	0.18	0.10	0.20	-0.05	0.09	0.20	0.17	0.03	0.88	-0.03	0.24	0.40	0.09	0.35
CVPA11	0.29	0.16	0.20	-0.01	-0.02	0.13	0.07	0.08	0.09	0.10	0.21	0.08	0.06	0.18	0.05	0.01	0.91	-0.09	0.22	0.38	0.10	0.49
DIS1	0.00	-0.45	-0.25	-0.46	-0.45	-0.31	-0.45	-0.49	-0.25	-0.37	0.25	0.81	-0.49	-0.32	-0.38	-0.67	0.10	-0.21	-0.46	0.05	-0.27	0.07
DIS2	0.00	-0.29	-0.23	-0.15	-0.22	-0.33	-0.34	-0.39	-0.18	-0.28	0.14	0.73	-0.44	-0.24	-0.22	-0.43	0.14	-0.18	-0.44	0.05	-0.24	0.19
DIS3	-0.05	-0.27	-0.14	-0.25	-0.25	-0.14	-0.33	-0.23	-0.19	-0.17	0.14	0.67	-0.29	-0.12	-0.23	-0.41	0.07	-0.11	-0.30	-0.07	-0.08	0.10
DIS4	0.02	-0.38	-0.22	-0.31	-0.20	-0.12	-0.31	-0.22	-0.21	-0.24	0.23	0.71	-0.26	-0.07	-0.16	-0.44	-0.01	-0.17	-0.26	-0.01	-0.11	0.01
DIS5	-0.19	-0.42	-0.35	-0.41	-0.45	-0.38	-0.41	-0.45	-0.40	-0.35	0.34	0.82	-0.42	-0.33	-0.30	-0.59	-0.11	-0.12	-0.54	-0.12	-0.37	-0.11
DIS6	-0.01	-0.45	-0.31	-0.24	-0.47	-0.40	-0.30	-0.46	-0.17	-0.38	0.37	0.82	-0.44	-0.31	-0.20	-0.60	0.06	-0.17	-0.52	0.05	-0.37	0.10
DIS7	0.08	-0.41	-0.42	-0.30	-0.42	-0.29	-0.37	-0.40	-0.24	-0.35	0.30	0.72	-0.37	-0.25	-0.33	-0.53	0.06	-0.01	-0.45	-0.03	-0.24	0.15
EA1	0.18	0.78	0.47	0.52	0.57	0.40	0.57	0.56	0.53	0.66	-0.05	-0.43	0.53	0.58	0.41	0.44	0.25	-0.09	0.54	0.21	0.37	0.20
EA2	0.05	0.81	0.53	0.48	0.55	0.38	0.49	0.60	0.50	0.73	-0.22	-0.35	0.44	0.45	0.37	0.41	0.20	-0.03	0.54	0.22	0.45	0.08
EA3	0.01	0.82	0.49	0.53	0.62	0.29	0.51	0.52	0.57	0.74	-0.06	-0.30	0.55	0.56	0.30	0.49	0.18	-0.15	0.52	0.15	0.36	0.14
EA4	0.17	0.84	0.41	0.60	0.71	0.45	0.56	0.64	0.60	0.82	-0.22	-0.47	0.64	0.64	0.31	0.52	0.17	0.04	0.62	0.16	0.44	0.08
G1	0.20	0.69	0.53	0.68	0.54	0.52	1.00	0.70	0.72	0.65	-0.19	-0.48	0.65	0.49	0.63	0.60	0.11	0.03	0.76	0.24	0.47	0.19
G2	0.11	0.76	0.49	0.54	0.68	0.69	0.70	1.00	0.59	0.81	-0.36	-0.52	0.80	0.63	0.46	0.66	0.10	0.10	0.79	0.10	0.59	-0.07
HQ1_1	0.15	0.42	0.41	0.51	0.40	0.41	0.53	0.47	0.80	0.56	-0.21	-0.21	0.48	0.44	0.31	0.29	0.09	0.09	0.47	0.20	0.46	0.21
HQ1_2	-0.02	0.72	0.44	0.50	0.55	0.48	0.49	0.74	0.54	0.84	-0.21	-0.44	0.74	0.58	0.29	0.58	0.02	-0.02	0.61	0.03	0.45	-0.15
HQ2_1	0.08	0.40	0.36	0.48	0.35	0.29	0.54	0.40	0.77	0.50	-0.26	-0.17	0.51	0.35	0.31	0.42	0.00	0.19	0.37	0.23	0.27	0.14
HQ2_2	0.15	0.74	0.41	0.61	0.59	0.48	0.65	0.67	0.70	0.86	-0.33	-0.27	0.57	0.52	0.35	0.52	0.08	0.03	0.57	0.15	0.48	0.14
HQ3_1	0.21	0.47	0.39	0.52	0.26	0.32	0.52	0.44	0.74	0.54	-0.22	-0.26	0.44	0.30	0.29	0.35	0.20	0.03	0.54	0.28	0.30	0.13

Appendix 6.3

Item	Scale																					
	ACU	AES	AI	B1	B2	BI	G1	G2	HQ1	HQ2	NEG	PD	PE	POS	PQ1	PQ2	RES	SA	SAT	TRU	UC	VAL
HQ3_2	0.03	0.65	0.46	0.29	0.54	0.34	0.40	0.51	0.43	0.71	-0.14	-0.24	0.47	0.47	0.22	0.37	0.15	0.08	0.53	0.13	0.36	0.15
HQ4_1	0.23	0.67	0.42	0.82	0.54	0.39	0.68	0.55	0.85	0.66	-0.16	-0.34	0.59	0.51	0.44	0.40	0.14	0.03	0.56	0.23	0.36	0.18
HQ4_2	0.13	0.79	0.39	0.66	0.74	0.50	0.59	0.75	0.70	0.91	-0.28	-0.43	0.73	0.66	0.27	0.56	0.10	-0.03	0.64	0.08	0.44	0.04
NEG1	0.14	-0.16	-0.16	-0.06	-0.05	-0.07	0.01	-0.06	-0.05	-0.07	0.50	0.26	-0.07	0.00	-0.03	-0.23	0.09	-0.21	-0.11	-0.06	-0.13	0.17
NEG10	0.03	-0.09	-0.03	-0.06	-0.14	-0.17	-0.12	-0.17	-0.06	-0.15	0.65	0.14	-0.13	-0.03	-0.19	-0.26	0.24	-0.22	-0.16	0.01	-0.14	0.17
NEG2	0.05	-0.21	0.00	-0.19	-0.27	-0.25	-0.12	-0.26	-0.08	-0.21	0.52	0.10	-0.09	0.00	-0.01	-0.10	0.21	-0.07	-0.18	-0.05	-0.08	0.16
NEG3	-0.10	-0.05	-0.06	-0.03	-0.16	-0.04	-0.04	0.01	-0.08	-0.04	**0.47**	0.11	0.09	0.17	-0.12	-0.04	0.02	-0.08	0.00	-0.05	-0.03	0.02
NEG4	-0.13	-0.08	0.05	-0.14	-0.17	-0.11	-0.01	-0.14	-0.13	-0.11	0.69	0.08	0.02	0.10	-0.01	0.01	0.11	-0.14	-0.02	-0.01	-0.08	0.16
NEG5	0.16	-0.11	-0.17	-0.02	-0.14	-0.23	-0.10	-0.28	-0.14	-0.13	0.63	0.12	-0.11	0.09	-0.13	-0.08	0.13	-0.22	-0.11	-0.05	-0.09	0.23
NEG6	-0.06	-0.30	-0.21	-0.42	-0.41	-0.32	-0.26	-0.40	-0.40	-0.40	0.73	0.42	-0.39	-0.16	-0.13	-0.29	0.03	-0.24	-0.34	-0.07	-0.30	0.04
NEG7	-0.03	-0.01	-0.02	-0.07	-0.15	-0.15	-0.07	-0.18	-0.07	-0.09	0.73	0.17	-0.07	0.05	-0.18	-0.16	0.19	-0.16	-0.10	0.09	-0.25	0.14
NEG8	0.00	-0.12	-0.06	-0.20	-0.22	-0.15	-0.11	-0.22	-0.15	-0.16	0.69	0.27	-0.15	-0.06	-0.13	-0.21	0.13	-0.10	-0.15	-0.05	-0.11	0.18
NEG9	-0.01	0.08	-0.08	0.00	0.21	-0.01	-0.11	-0.03	-0.03	0.07	**0.42**	0.14	-0.02	0.23	-0.22	-0.08	0.04	-0.17	-0.07	-0.09	-0.06	0.20
PE1	0.09	0.69	0.50	0.57	0.65	0.60	0.66	0.81	0.67	0.78	-0.30	-0.53	0.95	0.71	0.40	0.70	0.03	0.18	0.77	0.13	0.49	-0.03
PE2	0.15	0.65	0.49	0.51	0.56	0.46	0.64	0.76	0.61	0.71	-0.23	-0.54	0.96	0.70	0.39	0.73	0.07	0.08	0.71	0.06	0.49	0.01
PE3	0.33	0.49	0.33	0.38	0.50	0.40	0.48	0.62	0.48	0.62	-0.13	-0.36	0.88	0.68	0.22	0.47	0.11	0.05	0.57	0.00	0.41	0.07
POS1	0.22	0.56	0.42	0.50	0.54	0.60	0.54	**0.73**	0.57	0.67	-0.26	-0.47	**0.81**	0.69	0.33	0.60	0.13	0.16	0.67	0.12	0.60	0.06
POS10	0.30	0.42	0.19	0.27	0.49	0.34	0.29	0.41	0.19	0.34	0.08	-0.26	0.41	0.75	0.13	0.30	0.23	0.03	0.43	0.02	0.38	0.26
POS2	0.23	0.57	0.33	0.52	0.61	0.42	0.48	0.54	0.58	0.62	-0.05	-0.20	0.69	0.86	0.16	0.45	0.17	-0.04	0.49	0.10	0.41	0.22
POS3	0.23	0.46	0.18	0.39	0.57	0.26	0.29	0.39	0.38	0.47	0.15	-0.15	0.49	0.74	0.05	0.30	0.21	-0.21	0.32	-0.01	0.23	0.22
POS4	0.21	0.40	0.41	0.27	0.45	0.33	0.38	0.46	0.42	0.51	-0.11	-0.31	0.66	0.78	0.11	0.42	0.11	0.04	0.48	0.10	0.48	0.16
POS5	0.15	0.57	0.36	0.35	0.48	0.31	0.37	0.46	0.36	0.52	0.13	-0.20	0.49	0.78	0.20	0.35	0.20	-0.24	0.42	0.12	0.39	0.14
POS6	0.13	0.23	0.38	0.12	0.21	0.18	0.27	0.27	0.23	0.25	0.17	-0.18	0.38	0.57	0.23	0.24	0.03	0.05	0.29	0.08	0.30	0.04
POS7	0.27	0.55	0.37	0.26	0.53	0.32	0.31	0.44	0.35	0.56	-0.06	-0.23	0.53	0.79	0.13	0.34	0.16	-0.09	0.47	0.13	0.45	0.20
POS8	0.15	0.46	0.23	0.22	0.43	0.27	0.18	0.34	0.23	0.44	0.14	-0.11	0.43	0.78	-0.03	0.25	0.18	-0.04	0.32	0.01	0.42	0.19
POS9	0.20	0.40	0.44	0.24	0.30	0.28	0.32	0.37	0.26	0.36	-0.09	-0.19	0.42	0.63	0.25	0.34	0.17	0.07	0.45	0.25	0.44	0.16
PQ1_1	0.06	0.37	0.37	0.41	0.19	0.27	0.56	0.38	0.44	0.27	-0.20	-0.30	0.25	0.13	0.82	0.42	0.01	0.00	0.44	0.14	0.32	0.02
PQ1_2	0.14	0.56	0.38	0.53	0.53	0.42	0.60	0.62	0.51	0.58	-0.23	-0.62	0.65	0.50	0.46	0.88	-0.01	0.12	0.62	0.06	0.42	0.07
PQ2_1	0.03	0.05	0.02	-0.01	-0.09	0.13	0.05	-0.02	-0.12	-0.10	-0.05	-0.14	-0.09	-0.06	**0.23**	0.08	-0.05	-0.11	0.03	0.10	-0.03	-0.13
PQ2_2	0.08	0.25	0.07	0.14	0.11	0.11	0.11	0.24	0.08	0.19	-0.21	-0.35	0.19	0.06	0.21	0.50	0.06	-0.05	0.21	-0.04	0.08	-0.14
PQ3_1	0.12	0.38	0.47	0.27	0.27	0.30	0.50	0.42	0.28	0.30	-0.17	-0.30	0.32	0.22	0.86	0.39	0.06	-0.07	0.49	0.19	0.31	-0.05
PQ3_2	0.04	0.51	0.41	0.30	0.53	0.47	0.37	0.57	0.30	0.51	-0.41	-0.63	0.57	0.45	0.35	0.82	-0.03	0.18	0.58	0.14	0.42	-0.13
PQ4_1	0.19	0.39	0.41	0.32	0.18	0.19	0.52	0.34	0.36	0.28	-0.07	-0.27	0.37	0.20	0.81	0.43	0.22	-0.04	0.36	0.14	0.19	0.11
PQ4_2	-0.04	0.56	0.32	0.40	0.42	0.26	0.58	0.47	0.38	0.48	-0.02	-0.52	0.56	0.39	0.45	0.79	0.01	-0.06	0.52	-0.03	0.25	-0.06
SA	0.15	-0.02	0.00	0.05	0.05	0.20	0.03	0.10	0.11	0.01	-0.27	-0.18	0.12	-0.03	-0.04	0.09	-0.01	1.00	0.13	0.04	0.18	0.02
SAT1	0.24	0.67	0.57	0.50	0.53	0.70	0.73	0.81	0.58	0.65	-0.26	-0.55	0.72	0.58	0.49	0.62	0.19	0.18	0.94	0.30	0.70	0.14
SAT2	0.16	0.66	0.47	0.45	0.55	0.57	0.66	0.69	0.51	0.64	-0.24	-0.50	0.66	0.59	0.46	0.63	0.18	0.12	0.90	0.17	0.56	0.06
SAT3	0.13	0.69	0.70	0.47	0.49	0.57	0.68	0.67	0.61	0.65	-0.24	-0.54	0.67	0.52	0.47	0.61	0.27	0.05	0.91	0.46	0.57	0.11
TRU1	0.10	0.18	0.48	0.19	-0.02	0.24	0.19	0.07	0.25	0.08	-0.03	-0.08	0.01	0.08	0.10	0.09	0.31	0.13	0.26	0.80	0.32	0.15
TRU2	0.16	0.12	0.40	0.05	-0.04	0.33	0.21	0.06	0.17	0.00	-0.05	-0.01	-0.02	0.07	0.18	-0.01	0.34	0.03	0.29	0.91	0.30	0.20
TRU3	0.24	0.13	0.43	0.08	0.06	0.32	0.15	0.04	0.22	0.06	-0.04	0.02	0.05	0.11	0.06	-0.01	0.37	-0.01	0.27	0.91	0.29	0.32
TRU4	0.14	0.29	0.50	0.23	0.14	0.27	0.27	0.14	0.35	0.21	-0.06	0.02	0.19	0.17	0.29	0.12	0.38	-0.01	0.34	0.84	0.27	0.21
UC1	0.07	0.32	0.43	0.19	0.17	0.42	0.31	0.31	0.26	0.27	-0.22	-0.28	0.21	0.24	0.22	0.25	0.16	0.24	0.53	0.33	0.72	0.17

Appendix 6.3

Item	Scale																				
	ACU	AES	AI	B1	B2	BI	G1	G2	HQ1	HQ2	NEG	PD	PE	POS	PQ1	PQ2	RES	SA	SAT	TRU	UC	VAL
UC2	0.13	0.39	0.47	0.22	0.23	0.46	0.28	0.42	0.28	0.36	-0.24	-0.28	0.30	0.37	0.23	0.30	0.13	0.23	0.50	0.36	0.87	0.21
UC3	0.23	0.55	0.50	0.38	0.42	0.42	0.49	0.59	0.46	0.55	-0.14	-0.26	0.60	0.65	0.31	0.43	0.17	0.02	0.55	0.16	0.78	0.17

Note. Grey background indicates the loadings of items belonging to a particular scale. Bold numbers indicate loadings smaller than .50 and cross-loadings exceeding loadings. ACU: CVPA acumen. AES: aesthetics. AI: adequacy of information. B1: beauty before use. B2: beauty after use. G1: goodness before use. G2: goodness after use. HQ1: hedonic quality before use. HQ2: hedonic quality after use. NEG: negative affect. PD: perceived disorientation. PE: perceived enjoyment. POS: positive affect. PQ1: pragmatic quality before use. PQ2: pragmatic quality after use. RES: CVPA response. SA: spatial ability. SAT: satisfaction. TRU: trust. UC: usefulness of content. VAL: CVPA value. CA: classical aesthetics. EA: expressive aesthetics.

Appendix 7.1. A summary of the three studies presented in the current thesis

Source	Artefact	Task	Measures of experience	Person measures	Artefact measures	Task/context measures	Objective measures	Interaction outcomes
Study 1	Gazette Live news site	Free browsing under think-aloud instructions	Intensity of flow Perceived enjoyment Verbal protocols {D}	N/A	Pragmatic quality Hedonic quality	Level of adoption [novice/expert]	N/A	Beauty Goodness
Study 2	A news site of the participants' own choice	Free browsing	Perceived enjoyment Need fulfilment (5 selected needs) Affect (positive and negative)	N/A	Pragmatic quality Hedonic quality Usefulness of content Adequacy of information Accessibility Perceived disorientation Perceived aesthetics Perceived user-interface design	N/A	N/A	Beauty Goodness Intention to use
Study 3	BBC and NZH news sites	Observation (screenshots) and free browsing	Positive affect Perceived enjoyment	Centrality of visual product aesthetics Spatial ability	Pragmatic quality {B,A} Hedonic quality {B,A} Usefulness of content Adequacy of information Trust Perceived aesthetics Perceived disorientation	Level of adoption [novice/expert]	Number of pages visited Number of homepage revisits	Beauty {B,A} Goodness {B,A} Intention to use Satisfaction

Note. Conditions of experimental manipulations are presented in square brackets. Letters in curly brackets indicate times of measurement (B: before interaction, D: during interaction and A: after interaction). Measurement after interaction is treated as default; therefore, it is only indicated when another time of measurement is also present for a particular measure.